THE GREAT COURSES

Topic
Science
& Mathematics

Subtopic
Physics

Black Holes, Tides, and Curved Spacetime: Understanding Gravity

Course Guidebook

Professor Benjamin Schumacher
Kenyon College

PUBLISHED BY:

THE GREAT COURSES
Corporate Headquarters
4840 Westfields Boulevard, Suite 500
Chantilly, Virginia 20151-2299
Phone: 1-800-832-2412
Fax: 703-378-3819
www.thegreatcourses.com

Printed in the United States of America

Benjamin Schumacher, Ph.D.

Professor of Physics
Kenyon College

Professor Benjamin Schumacher is Professor of Physics at Kenyon College, where he has taught for 25 years. He received his B.A. from Hendrix College and his Ph.D. in Theoretical Physics from The University of Texas at Austin, where he was the last doctoral student of John Archibald Wheeler.

Professor Schumacher is the author of numerous scientific papers and books, including *Physics in Spacetime: An Introduction to Special Relativity*, and is the coauthor of *Quantum Processes, Systems, and Information*. As one of the founders of quantum information theory, Professor Schumacher introduced the term "qubit," invented quantum data compression (also known as Schumacher compression), and established several fundamental results about the information capacity of quantum systems. For his contributions, he won the 2002 International Quantum Communication Award, the premier international prize in the field, and was named a Fellow of the American Physical Society. Besides quantum information theory, he has done physics research on black holes, thermodynamics, and statistical mechanics.

Professor Schumacher has spent sabbaticals working at Los Alamos National Laboratory and the Institute for Quantum Information at the California Institute of Technology, where he was a Moore Distinguished Scholar. He also has done research at the Isaac Newton Institute for Mathematical Sciences of the University of Cambridge, the Santa Fe Institute, Perimeter Institute, The University of New Mexico, the University of Montreal, the University of Innsbruck, and The University of Queensland.

At Kenyon College, Professor Schumacher teaches physics primarily, but he also regularly ventures into astronomy, mathematics, scientific computing, and the humanities.

For The Great Courses, Professor Schumacher also has taught *Quantum Mechanics: The Physics of the Microscopic World* and *Impossible: Physics beyond the Edge*. ■

Table of Contents

INTRODUCTION

LECTURE GUIDES

Table of Contents

Table of Contents

Black Holes, Tides, and Curved Spacetime: Understanding Gravity

Scope:

Gravity is both the most familiar and the strangest of all the forces of nature. Though it is extraordinarily weak—it takes a whole planet to exert a fairly modest pull on a falling apple—it can act over astronomical distances. This is why gravity is the dominant force shaping the cosmos on the largest scale.

The study of gravity unites everyday phenomena with marvels of astrophysics. Galileo Galilei discovered the basic principles of motion and gravity on Earth, while Johannes Kepler worked out the laws of planetary motion around the Sun. These insights were incorporated by Isaac Newton in his magnificent system of mechanics, the science of force and motion. It was Newton who recognized that gravity is a universal force by which every mass attracts every other mass, everywhere.

Newton's discoveries became the foundation for all of physics for the next 200 years. Newtonian gravity accounts for the orbits of planets, moons, asteroids, and comets. It provides the basic rules for spacecraft navigation. It explains ocean tides on Earth and more spectacular tidal effects elsewhere in the universe. Gravity shows how the orbit of one body can be altered by the gravitational pull of another—an effect that allows us to "see the unseen" and discover new planets in our own solar system and beyond.

When three or more bodies orbit one another, their motions can be hard to predict. Orbital resonance can lead to surprising and complex results. Nevertheless, there are simple laws that apply even to the motions of millions of bodies, as in an entire galaxy. These gravitational laws enable us to detect vast amounts of "dark matter," an unseen mass whose gravity helps to hold the galaxies together. Closer to home, within such dense bodies as the Earth or the Sun, there is a balance between inward gravity and outward

pressure forces. This balance explains why a balloon can rise into the air—and it determines the structure and ultimate fate of a star.

It was Albert Einstein who launched the second revolution in gravitational physics. He began with two insights of Galileo: Absolute motion cannot be detected, and all objects fall freely with the same acceleration. From these principles, Einstein built his special and general theories of relativity—the first, a new way of thinking about space and time (or "spacetime"); the second, a startling new theory of gravity. For Einstein, gravity is not really a force but an expression of the curvature of spacetime. This curvature is, in turn, linked, via the tidal effect, to the density of matter. Einstein's new theory of gravity was summed up by physicist John Wheeler this way: "Spacetime tells matter how to move. Matter tells spacetime how to curve."

The implications of Einstein's theory of gravity are astonishing. Gravity affects everything, even time itself. A clock placed deep in a gravitational field runs slowly. Gravity also shifts the frequency of light waves and bends their path through space. It was Arthur Eddington's observation of light deflection by the Sun that led physicists to accept Einstein's ideas; nowadays, we can use gravity as a "lens" to probe the distant reaches of space.

A spinning planet can twist spacetime in its vicinity, and alterations in the curvature of spacetime "ripple" through the universe in the form of gravitational waves. Among the most remarkable implications of Einstein's theory are black holes, objects so massive and compact that nothing can escape the warped spacetime surrounding them. Yet black holes are not just exotic, theoretical creatures. They really exist. We have observed black holes throughout the universe, including a few that are billions of times more massive than our Sun.

Gravity governs the expansion of our universe and helps unravel many of the puzzles of the early days after the big bang. A kind of cosmic antigravity, called "dark matter" by cosmologists, causes the cosmic expansion to accelerate.

One of the strangest aspects of gravity is its paradoxical relation to entropy. The law of increasing entropy usually causes matter to disperse and become more uniform, but in the presence of gravity, exactly the opposite can occur. The most extreme example of this is in a black hole, which has an entropy related to the area of its event horizon. This connection between gravity and entropy may give us a clue to the greatest unsolved mystery of theoretical physics: how to reconcile Einstein's spacetime theory of gravity with quantum mechanics. ■

The Strangest Force

Lecture 1

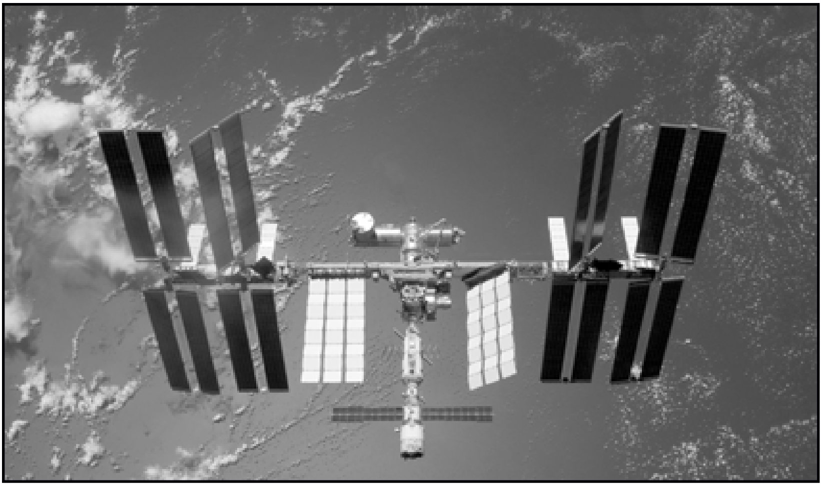
After seeing an apple fall in an orchard, Isaac Newton had a profound insight: that the same laws govern motion both on Earth and in the heavens and the same force that causes an apple to fall also steers the Moon in its orbital path. The idea that both the apple and the Moon are examples of gravity in action was a significant step forward in human thought and is a key aspect of our course. Gravity is at the heart of phenomena that are near at hand, on the human scale, everyday, and intuitive, as well as phenomena that are far off, at an astronomical scale, exotic, and surprising. It is a fundamental link between the commonplace and the cosmic.

Cosmic Collisions

- A supercomputer simulation shows the physics of one of the most stupendous events in the universe: the collision of two spiral galaxies. Each galaxy contains 100 billion stars, plus clouds of gas and dust. In the collision, the disk structure of each galaxy is completely disrupted. Huge arms of stars and gas are flung into space. Eventually, the crowded nuclei of the galaxies merge together and form a larger elliptical galaxy. This process might take 2 billion years.

- Galaxies are in the process of colliding all over the universe. Our own Milky Way galaxy has probably undergone collisions with smaller galaxies and is predicted to collide with its nearest large neighbor, the galaxy M31 in Andromeda, a few billion years from now.

- The most important force shaping such titanic events is gravity. Thus, from the simple fall of an apple toward the Earth to the complex dance of huge gas clouds and billions of stars in a galaxy collision, we are looking at the same thing: gravitational physics.



No place in the universe is beyond the reach of gravity; the strength of the Earth's gravity at the distance of the International Space Station is almost 90% of its strength on the ground.

Essential Characteristics of Gravity

- Gravity is a long-range force. This was Newton's original idea: that the gravitational pull of the Earth extends to the apple and out to the Moon. The Sun's gravity steers the planets in their orbits. Gravity holds each galaxy together, and it even reaches from one galaxy to another. Gravity gets weaker with distance, but there is no upper limit to its range.

- Gravitational forces are almost unimaginably weak.
    - A falling apple accelerates toward the Earth at about 10 meters per second per second (m/s^2), an acceleration of 1 *g*. An apple falling from your hand takes only about 0.5 seconds (s) to reach the ground. If you drop it from higher up, it might fall freely for just 3 s, but it is moving as fast as a car on a freeway. That's a respectable acceleration!

- But that acceleration is the result of the gravitational pull of the entire Earth—all 6 billion trillion tons of it. All that mass, acting together, produces an acceleration of only 1 *g*.

- Gravity is an attractive force. Given any two objects, the force of gravity pulls them toward each other. But this does not mean that they will actually get closer together.
  - For example, the Earth's gravity bends the Moon's path toward the Earth, but this just keeps the Moon curving in its orbit.

  - The attraction of gravity is mutual. Just as the apple is attracted to the Earth, the Earth is also attracted to the apple, but the acceleration of the Earth is undetectable.

  - Gravity works the same way for anything. There are no objects that are repelled from the Earth and fall upward.

  - This characteristic is also true in the collision of two galaxies. The streamers of stars and gas that are flung out into space are not actually repelled. It is simply that the gravitational attraction of one galaxy is stronger in some places—the near side of the other galaxy—and weaker in other places—the far side of the other galaxy. This difference in gravitational pull is called the tidal effect, and this is what disrupts the structure of each galaxy.

- Finally, gravity is a universal force. Newton realized that the Earth attracts not just the apple but also the Moon. In fact, each part of the Earth attracts each part of the apple and each part of the Moon. Newton called this the law of universal gravitation. Every two pieces of matter in the universe attract each other.
  - We see this at work in the collision of two galaxies. There are hundreds of billions of stars in the galaxies (not to mention gas clouds and so forth). Each star exerts a gravitational attraction on each other star.

- For this reason, such interactions are incredibly complex, requiring powerful supercomputers to understand the details.

Basic Forces of Nature

- We know of four fundamental forces that appear to account for everything we see: the short-range strong and weak nuclear forces and the long-range electromagnetic and gravitational forces.

- The strong nuclear force is a force between quarks, the basic constituents of protons, neutrons, mesons, and so on. This force extends over only about $10^{(-15)}$ m. The strong nuclear force holds quarks together within particles and holds atomic nuclei together. In that sense, it affects the motions of particles: It keeps them from moving apart.

- The weak nuclear force affects not only things made of quarks but also electrons and neutrinos. It is roughly 1 million times weaker than the strong force, and its range is only 1/1000 as far: $10^{(-18)}$ m. Thus, we see its effects only in particle transformations. The weak nuclear force, for example, is responsible for some kinds of radioactivity.

- The electromagnetic force involves two closely related phenomena: electric forces and magnetic forces. The electric force is a force between static electric charges. Objects can have positive or negative charges—zero is also possible—and like charges repel each other, while opposite charges attract. The magnetic force affects magnets and wires carrying electric current—electromagnets. About 200 years ago, physicists realized that both were different aspects of the same thing: interactions among electric charges, whether moving or at rest.
  - Like gravity, the electromagnetic force is long range. Its strength diminishes with distance, but it can exert some influence even from very far away. The electromagnetic force is moderately strong—100 times weaker than the strong nuclear force but thousands of times stronger than the weak nuclear force. It affects only particles that have electric charge—quarks and electrons—but it leaves neutrinos alone.

- Electromagnetic forces keep the electrons in an atom bound to the nucleus. This force is responsible for the short-range forces between molecules that hold solid objects together.

- Gravity, unlike the two nuclear forces, is long range; it can act over great distances. Unlike the electromagnetic force, it is always attractive. Unlike any of the other forces, gravity is universal, and it is weaker—much weaker—than any of the other forces.

Comparing Gravity and the Electromagnetic Force

- As we said, electric charges of the same sign repel each other. Of course, two electrons also attract each other by their mutual gravitation. Which is stronger—electric repulsion or gravitational attraction?

- Let *Fg* stand for the gravitational attraction of two electrons and *Fe* stand for their electric repulsion. What is the ratio of *Fg* to *Fe*? Conveniently, this ratio is the same no matter how far apart the electrons are, but it is an astonishingly small number: $Fg/Fe = 2.4 \times 10^{(-43)}$.

- Despite its weakness, gravity makes a significant difference in the universe. Electromagnetic forces are much stronger than gravity, but those forces tend to cancel out. Gravitational forces are tiny, but they are always attractive. All masses attract each other gravitationally, and the greater the mass, the greater the attraction. That means that for really large objects in the universe, gravity is by far the most important force.

- Consider the solid bodies orbiting our Sun: the planets and asteroids. What forces hold each of these bodies together? Basically, there are two candidates for the job: (1) intermolecular forces, which are short-range forces that reach from one molecule to the next, and (2) gravity, by which all parts of a body exert mutual attraction on each other. This is a much longer-range force, but as we have seen, it is very weak.

- A small piece of rock is held together by intermolecular forces. It's easy to see that gravity is insignificant here. If you create a thin crack all the way through the stone—a gap too wide for the intermolecular forces to reach across but no obstacle for gravity—then the rock doesn't hold together. It easily comes apart.

- If you want to vaporize the rock and disperse its molecules all over the universe, you will need to put in a good deal of energy—what physicists call "binding energy." Some of the energy that you put in will go into turning the solid rock material into vapor—that is, to overcoming the intermolecular forces—and some of the energy will go into dispersing the molecules—that is, to overcoming any gravitational attraction they have for each other.

- In the case of the rock, essentially 100% of the binding energy is due to the intermolecular forces, and almost none is due to gravity.

- Suppose you want to vaporize Mars and disperse its molecules all over the universe. How is the binding energy apportioned? In other words, which would take more energy, the vaporizing or the dispersing? It would take a huge amount of energy to vaporize a whole planet, but in fact, it would take several times as much energy to overcome gravitational attraction and disperse the molecules all over the universe. The binding energy of Mars is almost entirely due to gravity. Gravity is the main force that holds together a planet.

- Objects that are held together by short-range intermolecular forces have irregular shapes; each part is linked only to the parts immediately adjacent to it, and there is no tendency to smooth out irregularities. But a solid body held together by gravity tends to pull itself into a spherical shape. Each part of the body is attracted to every other part, even to pieces on the other side, and a sphere is the shape that minimizes the distances between all the pieces.

- Gravity is weak for small-scale objects, but on the largest scales, it is the most important force. Thus, in astrophysics, gravity is the most important force shaping the structure and motions of the universe. And in cosmology—the physics of the universe as a whole—gravity is king!

Suggested Reading

Feynman, *The Character of Physical Law*, chapters 1–2.

Questions to Consider

1. Suppose human civilization moves away from the surface of the Earth and establishes itself in deep space, where gravitational forces are negligible. How would the lack of gravity affect the activities of daily life, agriculture, architecture, the arts, and athletic competitions—to mention only the As?

2. Some physicists have speculated about the existence of a "fifth force" that would be nearly as weak as gravity but act only over a short range. Such a force would be extremely difficult to detect. Why? (As of this writing, no experimental search has found such a force.)

Free Fall and Inertia

Lecture 2

In the last lecture, we began our exploration of gravity with Isaac Newton and the famous story of the apple. But some of the most important insights about gravity and mechanics—the science of force and motion—actually predate Newton's work by more than half a century. Indeed, these basic discoveries were the inspiration not only for Newton but also for Einstein. If Newton was the father of mechanics, then the grandfather of mechanics was Galileo Galilei. In this lecture, we'll look at Galileo's three epoch-making discoveries about motion—the principle of inertia, the law of free fall, and the principle of relativity—which became the foundation of everything that came later in the science of gravity.

Galileo Galilei

- Galileo Galilei was born in 1564 and grew up in Pisa. He was the son of a famous musician and composer, Vincenzio Galilei. From his father, Galileo inherited a bent for mathematics, skill in intellectual controversy, and a penchant for experiment.

- Galileo studied mathematics and natural philosophy at the University of Pisa. He proved to be a shrewd observer of the natural world and thought deeply about the science of motion. In his 20s, Galileo became a university professor.

- Throughout his life, Galileo was interested in both theory and practicalities. He spent time with craftsmen, architects, and engineers and was something of a tinkerer. These interests gave him a sound practical knowledge of how things work, such as lenses, machines, ships, and so on, and encouraged his impulse to investigate the world through careful experiment and observation.

- Today, Galileo is famous as the inventor of the astronomical telescope, with which he made revolutionary discoveries about the

heavens. Just as important were his discoveries about motion, which served as the foundation of everything that came later in physics.

Principle of Inertia

- Aristotle believed that objects on Earth are "naturally" at rest. If something is moving, an outside influence must be causing the motion; remove the influence, and the object comes to rest. What needs explanation in this picture is velocity.

- Velocity is the rate of change in position. The formula for velocity is $v = \Delta x/\Delta t$, where x is position and t is time. The value for Δ can be either positive or negative.
  - For example, suppose at $t = 2$ s, you are at $x = 12$ m, and at $t = 4$ s, you are at $x = 6$ m.

  - Your velocity would be (–6 m)/(2 s) = –3 m/s. The minus sign indicates that you're moving in the "negative" direction.

- According to Aristotle, an object's velocity is naturally zero, but if we keep pushing it, it can continue to move at a speed that is not zero. Galileo believed that objects are not naturally at rest. In the absence of an outside influence, an object at rest does remain at rest ($v = 0$), but an object in motion also remains in motion at a constant velocity in a straight line. Inertia is the tendency of bodies to maintain a constant velocity, either a zero velocity or otherwise.

- According to Galileo, what requires a cause is not velocity but acceleration—the rate of change of velocity. The formula for acceleration is $a = \Delta v/\Delta t$. Any change in velocity counts as acceleration—speeding up, slowing down, or changing direction.

- Why do objects in everyday life slow down if they have this property of inertia? According to Galileo, the answer is friction, air resistance.
  - If you roll a ball on a flat table, it comes to rest. If you use a smoother ball and a smoother table or a heavy ball to reduce

the effect of air resistance, the ball will roll much farther before it comes to rest.

- If we imagine an ideal ball with no friction, it would roll forever, in agreement with the principle of inertia.

Law of Free Fall

- Once again, according to Aristotle, earthly objects naturally seek the center of the Earth, which is the center of the cosmos in the geocentric view. More massive objects have a stronger tendency to seek the center of the Earth; thus, more massive objects fall faster than less massive ones.

- For about 2000 years, this view was regarded as completely obvious. It was understood that an object increases its speed as it falls, but according to what rule? One answer was that an object that falls twice as far goes twice as fast.

- Galileo refuted these ideas. He discovered, first, that a freely falling object falls with a constant acceleration and, second, that all objects fall with exactly the same acceleration. We might call this pair of facts the law of free fall.

- With constant acceleration, as an object falls, its velocity increases with time. The formula for this is $v = at$, where v is the falling velocity; a is the acceleration, which remains the same; and t is the time since the object started falling. For freely falling objects near the Earth, a is 9.8 m/s (about 10 m/s).
- How far does the object fall? Remember, if the object falls faster each second, it covers more distance. The formula for the distance covered during constant acceleration is $x = \frac{1}{2}at^2$, where x is the distance fallen, a is the acceleration, and t is the time since the object started falling.
- The Museo Galileo has an interesting apparatus—a ball on an inclined track—to demonstrate Galileo's idea of accelerated

motion. Spaced at uneven intervals along the track are several hoops through which the ball rolls, and on each hoop is mounted a bell. The distance from the starting point to each hoop is a series of squares: 1 unit, 4 units, 9 units, 16 units, and so on. According to our acceleration equations, a rolling ball with a constant acceleration will ring the bells at equal time intervals. As the distance between the hoops widens, the speed of the ball also increases.

Galileo is well-known today for his many inventions, as well as his support of Nicolaus Copernicus's theory that the Earth and other planets move around the Sun.

Acceleration of Gravity

- Galileo's second discovery about gravity was that the acceleration of gravity is the same for all objects. According to legend, he made this discovery by dropping objects from the Leaning Tower of Pisa.

- Imagine that we drop two cannonballs, one weighing 100 pounds and one weighing 0.5 pounds. Our observation is that these hit the ground at almost exactly the same time. According to Galileo, the big cannonball leads by no more than a few inches. What conclusion can we draw?

- Aristotle would have said that the larger cannonball should fall 200 times faster, but that's wrong. Galileo's idea that the cannonballs should fall at the same rate is almost right. His explanation of the

tiny difference is that gravity is not the only force acting in this scenario. The cannonballs are not quite in free fall because of air resistance, which has a stronger effect on the smaller ball.

- Let's think about Felix Baumgartner and his record-setting jump from a balloon 39 kilometers (km) above the Earth.
    - After he first jumps, Baumgartner is in free fall. After 1 s, he is traveling at 10 m/s, having fallen 5 m. After 2 s, he is traveling at 20 m/s, having fallen 20 m. After 3 s, he is traveling at 30 m/s, having fallen 45 m. After 4 s, he is traveling at 40 m/s, having fallen 80 m.

    - These measurements agree with our equations for free fall, but after less than a minute, Baumgartner is no longer accelerating. In fact, he begins to slow down, because another force—air drag—becomes important. Once he is falling fast enough, the friction of the air prevents any further acceleration. This maximum speed is called his terminal velocity.

    - Terminal velocity depends on mass, shape, and air density. In Baumgartner's case, his maximum speed is more than 375 m/s, but as he falls through denser air, that terminal velocity decreases.

    - On Earth, at the surface, the terminal velocity of a falling human body is still fairly fast. On a planet with a much denser atmosphere, the terminal velocity might be much slower. With no atmosphere, a falling body continues to accelerate, and everything continues to accelerate at the same rate.

Principle of Relativity

- From Galileo's discoveries of the principle of inertia and the law of free fall came a third great discovery: the principle of relativity.

- In the Copernican theory, the Earth moves at tremendous speed, orbiting the Sun and rotating on its axis. One objection to this

theory is that if Earth were moving at such a high speed, we would feel that motion, but Galileo disputed that idea.

- Imagine climbing to the top of the mast of a ship at sea and dropping a ball down to the deck.
    - According to Aristotle, if the ship were at rest, the ball would fall straight down next to the mast. On a moving ship, as the ball falls, the ship would move forward, so the ball would hit the deck at a point behind the mast, toward the stern.

    - With this theory, if the Earth were moving at great speed, any object not actually tied to the Earth should lag behind it. The object would appear to be snatched backward. This is one argument the ancients made that the Earth itself had to be at rest.

- Galileo's analysis of this situation is different. The ball in this experiment has the property of inertia, like anything else. Initially, the ball is moving forward with the ship, and it continues to move forward as it falls. Seen from the sea, therefore, it follows a curved path downward. Seen from the ship, it still falls straight down next to the mast.
    - If we were to build a laboratory below deck on the ship and perform various experiments in it, we would not be able to tell from the experiments whether or not the ship was moving. Everything in the lab is initially moving along with the ship, and everything in the lab has the property of inertia, so it will continue moving in the same way.

    - Galileo envisioned many experiments for this lab, but every experiment would have the same results for either a ship at rest or a ship moving smoothly across the sea.

- Galileo's insight has some remarkable implications.
    - First, no experiment inside the laboratory can determine whether or not the laboratory is moving or at rest. Thus, no experiment on Earth can detect uniform motion of the Earth through space. In effect, we are all living in Galileo's ship.

- Second, absolute motion cannot be observed; only relative motion can be observed. Hence, this insight came to be called the principle of relativity.

Suggested Reading

Galileo, *Dialogue Concerning the Two Chief World Systems*. Galileo's discoveries about motion are described during the "Second Day" of the dialogue.

Schumacher, *The Pasadena Rule*.

Questions to Consider

1. What is the logical connection between the principles of inertia and relativity, both discovered by Galileo? To illuminate this question, suppose there were some object for which the law of inertia did not hold, that is, an object that was "naturally at rest," as Aristotle supposed. How would such an object behave in Galileo's shipboard laboratory?

2. Consider two objects, A and B. A falls twice as far as B, so it hits the ground at a higher speed. Some natural philosophers before Galileo believed that the object falling twice as far would be moving with twice the speed when it hit. Is this true?

Problem

How important was air resistance in Felix Baumgartner's long skydive from 39,000 m? Here is an interesting calculation: Show that, at a constant acceleration of 10 m/s^2, Baumgartner would fall to the ground in less than 90 seconds. (You should use Galileo's relation between distance, acceleration, and time.)

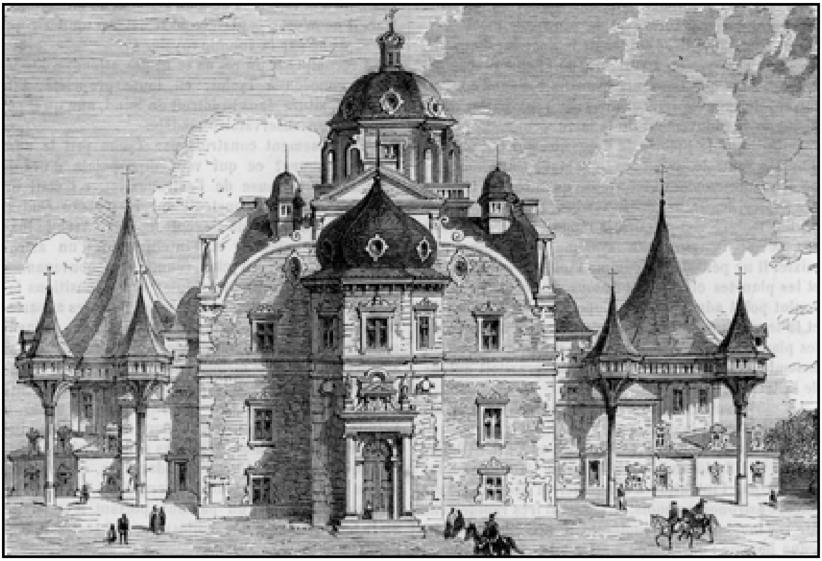
Revolution in the Heavens

Lecture 3

In the last lecture, we looked at three discoveries of Galileo: the principle of inertia, the law of free fall, and the principle of relativity. These tell us about motion and gravity for objects on Earth. Galileo, however, was concerned not just with the paths of projectiles but with the arrangement of the whole universe and our place in it. He invented the astronomical telescope, and with it, he made discoveries that changed our understanding of the universe. That revolution in the heavens is part of the story of gravity, too. In this lecture, we'll approach that story through the eyes of Galileo's contemporary Johannes Kepler, a great German mathematician, physicist, and astronomer.

Johannes Kepler and Tycho Brahe

- Johannes Kepler was a master of mathematics and had one other vital quality that allowed him to become a scientific pioneer: He was open to changing his mind. In his university days, Kepler studied astronomy and believed in the Copernican theory of the universe.
    - For most of history up to Kepler's day, almost everyone believed that the Earth was at the center of the universe, with the Sun, the Moon, and the planets moving around it.
    - This naïve idea was developed into a sophisticated mathematical system by ancient thinkers, most notably Claudius Ptolemy.
    - In Ptolemy's system, a planet moved on a small circle called an epicycle, and the epicycle itself moved on a large circle called a deferent, all moving around the Earth. The combination of these motions more or less explained how a planet moves in the sky.
- In the early 1500s, Nicolaus Copernicus proposed a radical theory: that the Sun, not the Earth, is at rest. The Earth is a planet that moves around the Sun with the other planets. It also rotates on its



Tycho Brahe recorded precise astronomical observations from his observatory, Uraniborg, the first great scientific research laboratory.

axis. According to Copernicus, the complicated motions of the other planets in the sky are due both to their own motion in space and the way our point of view moves along with the Earth. Copernicus fully explained his theory in his book *On the Revolutions of the Heavenly Spheres*, published in 1543.

- Kepler believed that the Copernican theory was a more elegant way to explain the motions of the heavens than the Ptolemaic theory. He also had some rather strange geometrical ideas about the Copernican system, including the notion that the sizes of the planetary orbits were somehow related to the properties of geometric solids. He published these ideas in a book, *Cosmographical Mysteries*, that caught the attention of Tycho Brahe, the greatest observational astronomer of his day.

- Tycho was a Danish nobleman who was interested in astronomy and dissatisfied with the then-current state of astronomical theories.

For 20 years, he and his assistants had created an exacting record of astronomical observations, and he knew that both Ptolemy's and Copernicus's theories were frequently off in their predictions of the motions of planets by as much as 5°. Still, like Ptolemy, he believed that the Earth was at rest in the center of the cosmos.

- Tycho brought Kepler to work for him and assigned him the task of explaining the motions of Mars. Shortly thereafter, Tycho died; Kepler took custody of all his observational records and continued to work on the problem of Mars.

The Motion of the Planets

- The actual Copernican theory is not quite as simple as we've described. In this theory, the shape of the orbit of a planet is a circle, but that circle is not centered on the Sun. To make the theory more accurate, Copernicus had to shift the orbits slightly.

- Kepler asked why a planet should move around a point in empty space. The path of a planet is not just an arbitrary curve. A planet moves through space on a certain path for some physical cause, and that physical cause, Kepler said, must be the Sun. The Sun must govern the laws of planetary motion.

- In his work on the problem of the orbit of Mars, Kepler asked two questions: What is the shape of that orbit, and what law governs the planet's motion around its orbit?

- To answer these questions, Kepler had to take into account the fact that all his observations were made from the Earth, which is itself moving. He had 20 years of observations from Tycho, from which he could correlate the motions of Mars and Earth. In almost 1000 pages of incredibly complex calculations—and from an infinite number of possible curves for the orbit—Kepler finally solved the problem, not just for Mars but for all the planets.

- In 1609, 66 years after Copernicus first published his theory, Kepler announced two laws of planetary motion. The first of these is the

law of ellipses: The orbit of each planet is an ellipse with the Sun at one focus.

  - Most planetary orbits are ellipses that are not too far from being circles.

  - To measure the size of an ellipse, it's natural to think of the major axis, the long axis of the ellipse. The semimajor axis is half of the long axis, a length that's generally denoted *a*. This is, roughly speaking, the average distance of the planet from the Sun. With a circular orbit, *a* is equal to the radius of the orbit exactly.

- Next was the law of equal areas: A line from the Sun to the planet sweeps out equal areas in equal times. When the planet is closer to the Sun, it moves faster in its orbit. When the planet is farther from the Sun, it moves slower in its orbit so that the wedge shape that's traced out has equal areas in equal times. For a circular orbit, the speed would be constant. Notice that in both of these laws, the planet, its path, and its motion are connected to the Sun, not to an arbitrary point in space.

The Harmonic Law

- Until Kepler's discovery, almost everyone believed that planets move through space because they have a motivating force inside them. But Kepler said that the force that steers the planet comes from the body of the Sun.

- Kepler didn't know the nature of this steering force, but he understood that it is a physical influence acting from a real object—the Sun. He also believed that the same kind of force acting from the Earth might steer the Moon in its orbit around Earth.

- Kepler noticed a qualitative trend. Planets with smaller orbits, such as Mercury and Venus, move quicker. The length of their orbital periods—the time for one orbit—is short. Planets with larger orbits, such as Jupiter and Saturn, move slower; the length of their orbital periods is much longer. In other words, there seems to be

a mathematical relationship between the size of the orbit and the orbital period, P, but the exact relationship is difficult to pin down.

- On the horizontal axis of a graph, we can plot the semimajor axis of each planet's orbit. On the vertical axis, we plot the orbital period. Each planet is a point on the graph. The points lie on a curve that sweeps upward, but what is the shape of that curve? Kepler wrestled with this problem for years. Then, in 1614, he learned of a new mathematical tool, logarithms, that simplified his astronomical calculations.

- If we return to our graph and plot the logarithms of the semimajor axis and the orbital period, we see that all the points representing planets lie on a straight line, and the slope of that line is exactly 3/2. From this discovery, Kepler worked out the relationship between the semimajor axis and orbital period. In 1618, he published his discovery, known as the harmonic law.

- The harmonic law states that the cube of the semimajor axis of a planet's orbit is proportional to the square of its orbital period. The formula here is $a^3 = KP^2$, where K is a constant that has the same value for all the planets.
    - The value of the constant, K, depends on the units of measurement used to measure the semimajor axis (a) and the orbital period (P).

    - Measuring a in astronomical units (AUs) makes the calculation simple because an AU is the radius of the Earth's orbit. The Earth's orbit has a semimajor axis of 1 AU. All of our distance measurements, then, are based on the Earth's orbit.

    - Time—the orbital period—is also based on Earth's orbit. We measure it in years, the orbital period of the Earth.

    - Thus, for the Earth, the value of K is a^3/P^2, which is 1/1. K has a value of 1, and that value is the same for all the planets.

  - We can use this value to do calculations about the planets. For example, Jupiter's orbit around the Sun has a semimajor axis of 5.2 AU. We can use Kepler's harmonic law to determine its orbital period. Using algebra, we find that $P = \sqrt{a^3 / K}$. If a = 5.2 and K = 1, then P = 11.9. Jupiter completes its orbit in 11.9 Earth years.

- The harmonic law applies to all the planets. It shows that they are all steered in their orbits by the same force, which Kepler realized must lie in the Sun. Further, when Galileo turned his newly invented telescope to the heavens, he discovered that planets other than Earth, such as Jupiter, have their own moons that form their own planetary systems. In the years after Galileo and Kepler, other astronomers found that the moons of Jupiter follow Kepler's law.
  - The orbits of Jupiter's moons are ellipses, in this case, nearly perfect circles. A line from Jupiter to each moon sweeps out equal areas in equal times, and the orbits are related to each other by Kepler's harmonic law: $a^3 = KP^2$.

  - In the case of Jupiter's moons, the value of K is not the same as it is for the solar system. It's much smaller for Jupiter and its moons than it is for the planets that orbit the Sun.

- Kepler's laws, it turns out, are not just about planets and moons. Anything that flies freely through the solar system, such as a spacecraft, follows those laws. The reason those laws apply to spacecraft is that they are guided by the same force that guides the planets, the same force that Kepler guessed holds the Earth's Moon in its orbit and steers the moons of Jupiter.

- Kepler never discovered that force, but Isaac Newton, half a century later, realized the astonishing truth. The force that steers the planets and the moons is actually the most familiar force of all: gravity. With this discovery, Newton was able to finish the revolution in the heavens that Kepler began.

Suggested Reading

Kepler, "The Sun as the Source of Planetary Motions," in *The Tests of Time* (Dolling et al., eds.). In an excerpt from his 1609 book containing his first two laws of planetary motion, Kepler argues for his conviction that planetary motions must be governed by the Sun.

Questions to Consider

1. Use the astronomer's trick mentioned in the lecture—that the width of a little finger at arm's length is about 1°—to consider the accuracy of Tycho's astronomical measurements. His instruments could locate a star or planet to within 1/30°. Can you resolve an angle that small with your own unaided eye?

2. It is sometimes said that Kepler's third law of planetary motion, the harmonic law, "completed" his theory. Explain this claim. What ingredient was missing in the first two laws, and how did the third law supply it?

Problem

The semimajor axis of Saturn's orbit around the Sun is 9.6 AU. Find the length of Saturn's year (i.e., its orbital period about the Sun).

Universal Gravitation

Lecture 4

In the last lecture, we looked at three mathematical laws describing the orbits of the planets devised by Johannes Kepler: the law of ellipses, the law of equal areas, and the harmonic law. Kepler knew that some force steers the planets in their orbits and that this force originates in the Sun, but he never knew what that force was. Fifty years after Kepler's discoveries, Isaac Newton found that the mysterious force that steers the planets is, in fact, the most familiar force of all: gravity, the same force that pulls the apple down from the tree. In this lecture, we'll follow Newton as he arrives at this insight and a far-reaching new framework for physics.

Newton and the Apple

- In a version of the apple story told by Newton himself, the falling of the apple inspired in him a question: Why does the apple fall down, that is, toward the center of the Earth? Why that direction out of all possible directions it might move?

- Newton realized that the center of the Earth has no special apple-pulling power or, rather, that the power is shared by everything in the universe. He said that every part of the Earth pulls at the apple. The net result of all those forces is to pull the apple straight down toward the center of the Earth.

- In other words, Newton found that the force of gravity is a universal property of matter. Every piece of matter in the universe—however small, however distant—exerts a gravitational pull on every other piece. Of course, large pieces of matter exert a greater power of attraction and are pulled more strongly by other pieces. The Earth's gravity pulls people with a greater force than it pulls the apple. Newton called this idea the law of universal gravitation.

Calculating Gravitational Force

- Suppose we have two particles, particle 1 and particle 2. They have masses m_1 and m_2, respectively, and they are a distance r apart. According to Newton, they are attracted to each other with a force, F, equal to $\frac{Gm_1m_2}{r^2}$, where G is a universal constant of nature (Newton's constant).
  - An equal force pulls particle 2 toward particle 1. Note that the force of gravity depends on both masses. Larger masses exert greater attraction, and they experience greater attraction.

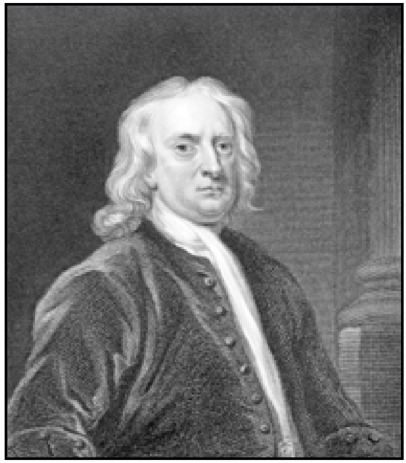
  - Note, too, that the equation calls for dividing by the square of the distance r. The greater the distance between the particles, the weaker the force of attraction, F. If the distance doubles, the force goes down by a factor of 2^2, or 4. If the distance is increased by a factor of 5, the force decreases by 5^2, or 25.

  - Finally, note that the constant, G, determines how strong the gravitational forces are. Newton himself never quite knew the value of his constant, and we'll look at it in greater detail in Lecture 5. For now, we'll say that the value of Newton's constant is: $G = 6.67 \times 10^{-11}$ m^3/kg s^2. The resulting value is a very small number, which tells us that gravity is a remarkably weak force.

- What about the gravitational force between large objects, such as stars, planets, and moons? Here, Newton discovered something amazing.
  - Consider the force of the whole Earth on an apple. The Earth is composed of zillions of tiny pieces of matter distributed through a huge volume of space. Some pieces are close by; others are far away, and each piece exerts its own pull on the apple.

  - Newton showed that as long as the Earth can be regarded as a sphere, the net result is exactly as if the entire mass of the Earth were located exactly at the center of the Earth.

- We can then use Newton's law of gravity to calculate the force exerted by the Earth on the apple. If M is the mass of the Earth, m is the mass of the apple, and R is the radius of the Earth, then an apple at the surface of the Earth is pulled with a force, F, equal to $\frac{GMm}{R^2}$. That's exactly as if the whole Earth were one massive particle located a distance R away.



© iStockphoto/Thinkstock.

It is impossible to exaggerate Newton's contributions to mathematics, mechanics, or optics, and for a century, physics was essentially a series of footnotes to Newton's work.

- What about the gravitational force between the Sun and the planets? Newton realized that the planets are steered in their paths by the Sun's gravity—a force that pulls each planet toward the Sun.
  - Kepler seems to have believed that the Sun's force pushed the planet along in its orbit, but Newton knew better. Galileo had shown that if no force were to act on a planet, then its property of inertia would carry it along in a straight line in space. The force is the thing that bends the path of the planet.

  - The force that deflects the planet from a straight-line path is a force that pushes toward the Sun. This is true anywhere along the planet's orbit. The path of the planet is always bending toward the Sun.

  - The force that steers a planet in a circular orbit is a centripetal force, a force that is directed toward the center of the circle. That force is gravity exerted by the Sun. This does not result in the planet getting nearer to the Sun in a circular orbit. It

stays at the same distance. As it follows a circular orbit, it continually turns toward the Sun, pulled in that direction by the Sun's gravity.

Newton's Laws of Motion

- To use his theory of gravity to understand motion, Newton also introduced three general laws of motion. The first of these is Galileo's law of inertia: In the absence of an external force, a body maintains a constant velocity; thus, a body at rest will tend to remain at rest, and a body in motion will tend to remain in motion in a straight line as long as there is no external force, which causes a change in velocity.

- Newton's second law is the equation of motion: $F = ma$, where F is the net force on a body, m is the mass of the body, and a is the resulting acceleration of the body.
    - The equation of motion contains some qualitative information. Suppose we apply a particular force to an object. Then, a small body with a small mass will have a large acceleration, but a large body with a large mass will have a small acceleration. Another way to think about this is as follows: Mass is a measure of the inertia of an object. It tells us how much it resists changes in its velocity.

    - If an object is not accelerating, then the net force on it must be zero, but that zero might be the result of several forces acting in different directions.

    - Finally, consider the units of measurement for the terms in the equation. Mass is typically measured in kg. Acceleration is the rate of change of velocity, which is measured in m/s. It's often written as m/s^2. The units of force must be kg × m/s^2. That combination is called 1 newton (N) of force. That means that the gravitational force on a 1-kg object is about 10 N.

- Consider a falling apple, accelerating downward due to the gravitational attraction of the Earth. The force of gravity is

$F = \frac{GMm}{R^2}$, where M is the mass of the Earth, m is the mass of the apple, and R is the radius of the Earth.

- Combining this with Newton's equation of motion, we get ma—the acceleration—equals $\frac{GMm}{R^2}$.
- Notice that the mass of the apple, m, appears twice in this equation. We can divide by m and get the acceleration of gravity: $a = \frac{GM}{R^2}$. This acceleration is designated g. In Lecture 2, we saw that g equals about 10 m/s^2.
- This acceleration depends on the mass and radius of the Earth but has nothing to do with the particular falling object; thus, Newton's theory gives us Galileo's law of free fall. All kinds of objects near the Earth's surface fall with exactly the same acceleration.
- The equation of gravitational acceleration allows us to figure out how strong gravity might be on another world, such as our own Moon or Jupiter.

- Newton's third law is the law of action and reaction. If object A exerts a force on object B, then object B exerts a force on A that is equal in strength and opposite in direction. If the Earth exerts a gravitational force on the apple, then the apple also exerts a force of the same size on the Earth. That equal gravitational force produces a much smaller acceleration for the Earth, which is much more massive than the apple, but the force is there. Everything exerts gravitational attraction on everything else.

- Newton applied his laws of motion and universal gravitation to the orbits of the planets. In the end, he was able to derive all of Kepler's laws. Further, Newton did not just prove Kepler's harmonic law relating the orbital period of a planet to the semimajor axis of its orbit, but he improved it.

- Kepler's equation was $a^3 = KP^2$, where K is the same constant for all the planets.
- Newton found that if a planet orbits the Sun with an orbital period P and a semimajor axis a, then $a^3 = \frac{G}{4\pi^2 MP^2}$, where M is the mass of the Sun. Whereas Kepler had an arbitrary constant in his proportionality law, Newton has an exact equation.
- This useful equation enables us to determine the masses of distant objects in space. We can observe P and a for some planet, and that allows us to figure out M, the mass of the Sun. We can determine the mass of Jupiter by looking at its moons according to the same laws.
- This technique is used for determining the masses of distant objects. Of course, the technique depends on two things: (1) measuring distances so that we can find the value of a for a distant orbit, and (2) knowing the value of Newton's constant, G. Before we learn more about Newtonian gravity, we'll look at how those two factors were determined.

Suggested Reading

Newton, *Philosophiae naturalis principia mathematica*, in *On the Shoulders of Giants* (Hawking, ed.). Though reading the whole of the *Principia* would be a considerable challenge, browsing the first two sections ("Definitions" and "Axioms") is well worth the time.

Stukeley, "Newton's Apple."

Questions to Consider

1. Newton's law of universal gravitation—that every mass exerts an attractive force on every other mass—was an amazing leap of imagination. What reasons could Newton have provided for adopting such a radical hypothesis?

2. Two planets have exactly the same total mass, but one is denser than the other one and, thus, has a smaller radius. Which one has a greater surface gravity?

Problem

The mass of the Earth is about 6×10^{24} kg, and the radius of the Earth is 6.4×10^{6} m (about 6400 km). From this information and the value of Newton's constant, G (6.67×10^{-11} m^3/kg s^2), calculate the acceleration due to gravity near the Earth's surface. Does the value you obtain agree with the value we gave in the last lecture (about 10 m/s^2)?

The Art of Experiment

Lecture 5

In the last lecture, we had our first look at Isaac Newton's magnificent theory of mechanics, including his three laws of motion and his law of universal gravitation. Newton showed that the familiar force of gravity, the force that makes the apple fall from the tree, is also the force that steers the planets and moons in their orbits. He was able to derive Kepler's laws of planetary motion as consequences of his theory. The same force—the same physics—applies both on Earth and in the heavens. In this lecture, we'll see whether his theory is right and how we can apply it to the real world.

The Problem with Newton's Theory

- Suppose an object is at a distance R from the center of a mass M. This mass might be an apple near the surface of the Earth or a planet far from the Sun. The object experiences a gravitational force that causes it to accelerate. We found that acceleration with the formula $a = \frac{GM}{R^2}$, where G is Newton's constant. G has a very small value, which means that gravity is a very weak force.

- Notice that this acceleration does not depend on the mass of the apple, or the planet, or any of the planet's properties. The same object would accelerate the same amount if it were at the same distance, R, from the same mass, M.

- Unfortunately, in 1666, when Newton first formulated his theory, none of the values for the variables or constants was known. To overcome this difficulty, 17th-century astronomers and physicists had to conduct the most careful observations and the most delicate experiments ever performed. It took more than two centuries for all these questions to be firmly settled.

Determining Distances

- Copernicus knew how to determine the ratios of distances with some fairly straightforward observations. He worked out that the orbit of Mars is about 1.52 times larger than the Earth's orbit. The radius of the Earth's orbit is 1 AU. The orbit of Mars, therefore, has a semimajor axis of 1.52 AU. We can measure all the orbits of the planets in AUs, but how big is our yardstick? How far is the Sun from the Earth? In the 17th century, no one knew that distance.

- At around the same time that Newton was working with his gravitational ideas, a French astronomer named Giovanni Cassini set up a clever observation to determine the size of the AU.
    - In the early fall of 1672, the planet Mars came comparatively close to the Earth. In fact, Mars was not only near the Earth, but it was also near the point on its orbit closest to the Sun, meaning that Mars would be especially close to the Earth.

    - Cassini set about determining how far away Mars was from the Earth in September of 1672; to do this, he had to make simultaneous measurements of the position of Mars in the sky from two different locations on Earth. He dispatched an assistant to a French colony on an island off South America to measure the position of Mars. He then combined his measurement and his assistant's, along with the exact distance between the two observation points and figured out the distance to Mars.

    - Once that single distance was known, all the other distances in the solar system followed. Cassini established the basic yardstick for all measurements of distance beyond the Earth.

    - Today, we know that the value of the AU, the size of the Earth's orbit, is 1.5×10^{11} m, or about 150 million km.

Determining Mass and the Gravitational Constant

- How do we determine the mass of the Earth or the mass of the Sun? How do we determine the gravitational constant, *G*? These two

questions are actually related. If we can answer one, we can answer both because near the surface of the Earth, the value of R is about 6400 km. That's the radius of the Earth. The acceleration of gravity is 10 m/s^2. If we knew the value of G, we could solve the equation and determine the mass, M.

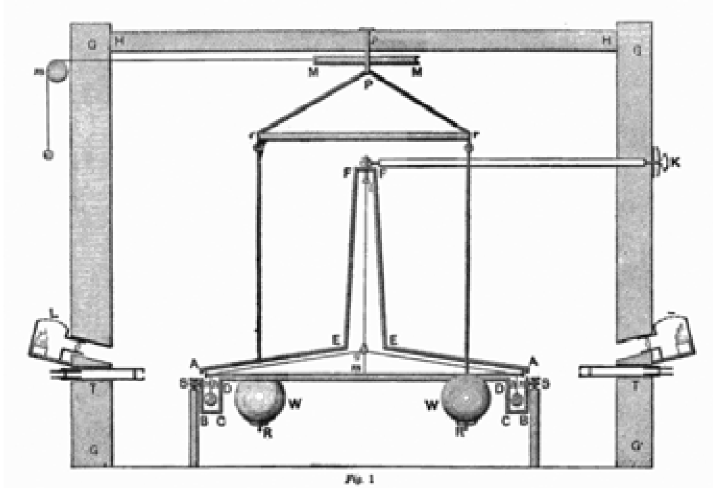
- Newton himself never learned either the value of G or the value of M for the Sun or the planets. He did know the value of the product GM, and that combination is what usually appears in the equations. Newton never found out just how strong the gravitational attraction is, and he never found out the mass of any astronomical object. In the end, he didn't have to; only the product GM mattered for his calculations.

- But that leaves a rather large unanswered question. Newton claimed that every object exerts an attractive force on every other object and the strength of this universal force is determined by the value of G—one of the fundamental parameters of nature. What would we have to do to measure G?
    - We would have to take a pair of known masses in the laboratory, put them a known distance apart, and then measure the extraordinarily tiny gravitational force between them.

    - How small is this force? If we have two 1-kg masses, 10 cm apart, the attractive force between them is only a few billionths of a newton.

- Incredibly, this experiment was done in 1798 by the brilliant and eccentric English scientist Henry Cavendish, using an apparatus called a torsional balance that looks similar to a hanging dumbbell.
    - Even a very tiny force can cause the dumbbell to turn a small amount. That motion causes the thin line to twist slightly.

    - We see a demonstration using brass shot puts. By measuring how much the gravitational force twists the line and conducting other experiments to determine how much force produces how much twist, we can determine the gravitational force.



With his torsional balance, Henry Cavendish essentially weighed the Earth; he designed a method for determining the mass of objects by observing their gravitational effects on other nearby objects.

- Modern experiments set the value of *G* at 6.67×10^{-11} $m^3/kg\ s^2$. Once we know this number, we can determine the mass of the Earth, as we described earlier, and that's what Cavendish did. The way we determine the mass of most things in the universe—moons, planets, stars, and so on—is by observing the gravitational effects of these objects on other nearby objects. Knowing the strength of the gravitational force—the value of Newton's constant *G*—we can deduce the mass.

Inertial Mass and Gravitational Mass

- In Newtonian mechanics, the mass of an object plays two different roles. The first is as a measure of the object's inertia, that is, its resistance to changes in velocity. This is the mass we find in Newton's second law, and it's called inertial mass.

- The object's mass also plays a role in the formula for the gravitational force on the object. That is, the mass of the object

determines how strongly the gravity of the Earth pulls at it. This is called gravitational mass.

- Inertial mass and gravitational mass are always exactly the same. This fact is sometimes called the principle of equivalence. Because of this principle, we can cancel the mass out of our equations. We find that everything in our laboratory falls with the same acceleration, just as Galileo observed.

- Why is inertial mass always equal to gravitational mass? Newton's theory doesn't appear to give us any real reason. It's just one of those things that happens to be true—or is it? Are inertial and gravitational mass always exactly the same for all objects, no matter what they're made of, no matter what the circumstances? Do all objects really fall with exactly the same acceleration?

- About 100 years after Newton, the French physicist Charles-Augustin de Coulomb determined the law governing the electric force between two charges. His equation for the force between charges Q_1 and Q_2 looks similar to Newton's law of gravity: $F = \frac{kQ_1Q_2}{R^2}$, where k is a constant expressing the strength of the electric force.
    - The electric force does not depend on the mass of an object but on the electric charge, which is an entirely independent property. Two particles with the same charge might have different masses. Two particles with the same mass might have different charges.

    - If we set up a large static charge somewhere and studied how other charges are accelerated by it, we would not find that they all accelerated in the same way. Some would be accelerated quickly; some, more slowly. Positive charges would be accelerated one way; negative charges, the opposite way. Neutral particles would not be accelerated at all.

- Gravity, on the other hand, is strange and unique. Everything is accelerated by gravitational forces in exactly the same way.

- Galileo did his experiment by dropping masses side by side, and he found that any differences in acceleration were very small. He attributed those differences to air resistance, friction, and so on.

- Newton did some better experiments using pendulums made of different materials. He found that the principle of equivalence held good to about 1 part in 1000.

- But such an important and unique feature of gravity needs a solid experimental basis. That basis was provided by a series of amazingly precise experiments, carried out between 1885 and 1910 by Baron Lorand Eotvos and his assistants.
    - Like Cavendish, Eotvos used a torsion balance. Eotvos dispensed with the large outside masses and made the two masses at the ends of the dumbbell out of different material, such as iron and brass or glass and wood.

    - He allowed the balance to come to rest and then watched it very closely. Of course, the gravitation of the Earth affects both ends according to their gravitational masses, and as the Earth turns on its axis, most masses move in circular paths around the Earth, and they resist that acceleration according to their inertial masses.

    - In Newton's theory, where gravitational and inertial masses are equivalent, these effects work out the same for the two masses, so there would be no net effect on the torsion balance. If the principle of equivalence were not exactly true—if the objects on the two ends have slightly different ratios between their gravitational and their inertial properties—then, as the Earth rotates, the two ends of the dumbbell would respond slightly differently.

    - Eotvos designed and built a series of increasingly sensitive torsional balances, and in the end, his experiments allowed him to conclude that the principle of equivalence is true to better

than 1 part in 100 million. Nowadays, with improvements in technology, we can push that precision to 1 part in 100 billion.

- The validity of the principle of equivalence means that Galileo was exactly right, as far as we can tell, when he said that everything falls at the same rate. This same principle, so strikingly confirmed by Eotvos, later became the cornerstone of Einstein's astonishing new theory of gravity.

Suggested Reading

Gamow, *Gravity*, chapter 2. Gamow's discussion is at a higher mathematical level than ours, but his discussion of the Cavendish experiment is useful.

Wheeler, *A Journey into Gravity and Spacetime*, chapter 2. This chapter includes a discussion of Eotvos's experiment on the principle of equivalence.

Questions to Consider

1. Before Cavendish's experiment, the value of the gravitational constant G was not known. Newton nevertheless knew that the Sun is more than 300,000 times more massive than the Earth. How could he know this fact? Would his reasoning have relied on the measurement of solar system distances by Cassini and Richer?

2. Eotvos's experiments with the torsion balance confirmed the principle of equivalence to fantastic precision. Could such precision be obtained by observing the free fall of dropped objects, à la Galileo? Why or why not?

Problem

Use the Newtonian equations from the previous lectures to calculate the acceleration of gravity due to a 7.25-kg shot put at a distance of 0.10 m from its center. At this acceleration, how far would an object "fall" in 1000 seconds (a bit more than a quarter of an hour)?

Escape Velocity, Energy, and Rotation

Lecture 6

In this lecture, we aim to get a better understanding of the way that forces, especially the force of gravity, affect motion. We will use two key concepts for describing how gravity affects motion: energy and angular momentum. Energy comes in several types, such as kinetic energy and potential energy. Angular momentum is a way of measuring the circular motion of a system. The useful thing about these two quantities is that they are conserved, that is, both the total energy and the angular momentum of a collection of masses interacting by gravity remains constant over time.

Kinetic Energy

- The idea behind kinetic energy (KE) is simple: Objects have energy simply because they are moving.
    - A car moving along a highway has both a mass, *m*, and a speed, *v*. Its KE depends on its mass and speed, that is: KE = $\frac{1}{2}mv^2$
    - Because the KE value depends on the mass, a larger car has more KE than a smaller car. Because the KE value also depends on doubling the velocity, a car has greater KE the faster it goes.

- We measure mass in kg and speed in m/s; KE is calculated in a unit called a joule (J). One J is the KE of a 2-kg object, for instance, a 2-liter (L) bottle of water, moving at 1 m/s, a moderate walking pace.

- The KE of an object is not constant. It might change when a force acts. If you toss an apple upward, gravity slows it down, so its KE decreases. If you drop the apple downward, gravity speeds it up, so its KE increases. If the apple is in a circular orbit around the Earth moving at a constant speed but continually changing direction, then its KE stays the same.

- Whenever a force produces a change in the KE of an object, physicists say that the force does work on the object. The term "work" here actually has a technical meaning in terms of forces and movements, but all we need to know for now is that the total work done on an object is the change in its KE.

Potential Energy

- Potential energy (PE) is always connected to some force; for us, this force is gravity. PE describes the potential of a force to do work. When gravity does work on something, the PE changes by the opposite amount. Although KE depends on how things are moving, PE depends on where they are. It is energy stored in the location of something.

- Again, consider dropping an apple. The PE is greater when the apple is higher up and lower when the apple is farther down. At the beginning, PE is greater, but KE is zero. But as the apple falls, its KE increases while its PE decreases. If we add up the two kinds of energy, we arrive at the total mechanical energy.

- As the apple falls, its mechanical energy stays constant; this is true whether the apple is going up or down or sideways in orbit around the Earth. As long as gravity is the only force acting on it, KE + PE stays the same.

- To determine a formula for PE due to gravity, we first need to answer a rather peculiar question. Consider the Earth and the apple as before. When should PE = 0?
    - It's easy to see when KE = 0. KE is energy of motion, and it's 0 when the object is at rest. But where should we put the apple so that it has 0 PE?

    - Our idea is to set PE = 0 where the gravitational force is also equal to 0. The gravitational force gets weaker with distance, so we will set PE = 0 for the apple when it is extremely far away, effectively at an infinite distance.

- Remember, we've said that the lower the apple is—the closer it is to the Earth—the lower its PE will be. Lower than zero means negative. The gravitational PE of the apple will always be a negative quantity. When the apple falls toward the Earth, its PE goes down by getting more negative, further below the zero point.

- Imagine two masses, m_1 and m_2, separated by a distance r between their centers. The equation for gravitational PE is $-\frac{Gm_1m_2}{r}$. As we said, PE is negative. Because r is in the denominator, PE is close to 0 when r is very large. The PE is about 0 for the two masses when they are far apart, and then it gets more negative as they approach each other.

Escape Speed

- When the only force acting is gravity, total mechanical energy stays constant; this fact allows us to figure out something quite interesting.
  - If you toss an apple, it goes upward, slowing until it reaches the top of its path, where it is momentarily at rest—KE = 0 at the top—then it falls back down. If you toss the apple a little faster, it goes a little higher. At the top of its trajectory, its KE = 0, which is the smallest possible value for KE.

  - In each case, the total energy is negative. PE for gravity is always negative, even at the top of the arc, and because KE = 0 there, the total energy is also negative.

- What if you throw the apple upward so fast that it goes far from the Earth? Out there, Earth's gravitational pull is much weaker, and the apple hardly slows down at all. At very great distances, the gravitational force is negligible, and the apple would fly away at a nearly constant speed. In other words, if you could launch the apple fast enough, it might never fall back to Earth, but how fast is fast enough?

To launch any object into space, a rocket must be used to accelerate the object to more than the Earth's escape speed, which is greater than 11 km/s.

- The escape speed is the minimum speed necessary for the apple or any object to completely escape Earth's gravity.
  - If the apple escapes Earth's gravity, then it will eventually reach a huge distance from the Earth. With such a separation between the Earth and the apple, PE is practically 0. Of course, KE cannot be less than 0 no matter how slowly the apple moves.

  - That means that for an apple that is escaping, the total KE + PE must be at least as great as 0. Therefore, the escape speed is the speed the apple needs so that KE + PE = 0.
- If we let m be the mass of the apple, M be the mass of the Earth, and r be the initial distance between the apple and the center of the Earth, we get the following formula for the escape speed, where

the first value is KE and the second is PE: $\frac{1}{2}mv^2 - \frac{GMm}{r} = 0$. We can then solve this equation for the speed, v, and doing so, we get: $v_{\text{escape}} = \sqrt{\frac{2GM}{r}}$.

- We've said that mechanical energy stays constant, but that's true only if gravity is the only force acting. When other forces act, we have to take into account other forms of energy.
  - When you throw the apple upward, where does the KE come from? The answer is chemical energy stored in your muscles.

  - Electric energy, nuclear energy, and so on can all be turned into mechanical energy. The total energy we add up is not only KE and PE but also all other types of energy.

  - Total energy is always conserved. Energy can change form or be transferred from one place to another, but it cannot be created or destroyed. That's the law of conservation of energy.

Thermal Energy

- If you drop the apple, as it falls, PE is turned into KE, but when it hits the ground, it bounces and rolls a bit and then comes to rest. Where did the energy go? It became thermal energy—heat.

- Objects have internal energy simply because they are warm. In fact, the temperature of an object is a measure of its heat energy content. Higher temperature means more thermal energy. After the apple comes to rest, both it and the ground have warmed up by a tiny amount. PE turns into KE, which then turns into heat.

- The KE of molecular motion is what we perceive as temperature. Thus, heat energy is not really a new kind of energy; it's actually familiar mechanical energy—the energy of the disordered, random motions of the molecules that make up objects.

Linear and Angular Momentum

- Like KE, the linear momentum of an object depends on its velocity and mass. The formula for momentum is simpler: $p = mv$. Momentum does depend on the direction an object is moving—what physicists call a vector quantity. This has both magnitude and direction. The momentum of a body has a certain value south, or up, or north by northwest, depending on the direction it is traveling.

- Total momentum, like energy, is conserved. This fact is derived from Newton's law of action and reaction.
  - Suppose object A exerts a force on object B. That force causes B to accelerate, to change its momentum. B is also exerting a force on A. That force is of the same size but opposite in direction and causes the momentum of A to change in the opposite way. The total momentum of A and B remains constant.

  - This suggests a different way of thinking about force: When one object exerts a force on another, the two objects are exchanging momentum. A force is a transfer of momentum.

- Angular momentum is a general idea, but it's easiest to think about in a simple case. Suppose we have a mass, m, that's moving in a circular path of radius r at a speed of v. The angular momentum of the mass (ℓ) is given by: $\ell = mrv$.

- Angular momentum has its own law of conservation. In the absence of external forces, the total angular momentum of a system of bodies stays constant. Parts of a system can exchange angular momentum with other parts of the system, but the total remains the same.

- Angular moment is a vector, like linear momentum, but its direction is somewhat surprising. As a mass goes around in a circle, it travels in different directions at different times: first north, then west, then south, then east. If angular momentum is constant,

then it can’t point in any of those directions. The one direction that does not change is the direction perpendicular to the circle—up.

- The conservation of angular momentum is an important principle to keep in mind as we explore the way gravity shapes the universe.

Suggested Reading

Einstein and Infeld, *The Evolution of Physics*, chapter 1. Einstein and Infeld do a good job of describing the essentials of Newtonian mechanics.

Gamow, *Gravity*, chapter 8. Gamow gives a slightly more mathematical account of the idea of gravitational potential energy and escape speed.

Questions to Consider

1. The next time you visit a fair or an amusement park, consider how energy, momentum, and angular momentum all play roles in various rides. The roller coaster is especially interesting, as are the various pendulum swing rides that have names ranging from Pirate Ship to Frisbee.

2. Is the total mechanical energy (KE + PE) of the solar system positive or negative?

3. We saw that the planet Saturn is approximately 100 times as massive as the Earth and 10 times the radius. This means that the force of gravity at Saturn’s surface is about the same as on Earth. Is the escape speed from Saturn greater than, about equal to, or less than the escape speed from Earth?

Stars in Their Courses—Orbital Mechanics

Lecture 7

In his epic of physics, the *Principia*, Newton imagined an artificial Earth satellite: a cannonball fired from a high mountain at a speed fast enough to allow it to follow a circular path around the planet. Newton used this thought experiment to explain his ideas about orbital motion. Of course, he wasn't really thinking about cannonballs but about the orbits of planets and moons. In this lecture, we will once again follow his lead to see how the concepts of mechanics, including energy and angular momentum, apply to orbital motion or anything moving around the Sun, the Earth, or anywhere in the universe.

Circular Orbit Speed

- As we saw earlier, Kepler described three laws of planetary motion: (1) the law of ellipses: the orbit of a planet is an ellipse with the Sun at one focus; (2) the law of equal areas: a line between the Sun and a planet sweeps out equal areas in equal times; and (3) the harmonic law: the cube of the semimajor axis of a planet's orbit is proportional to the square of its orbital period.
- Later, Newton explained Kepler's laws, which followed from his own laws of motion and universal gravitation. According to Newton, the Sun pulls on a planet with force (F) = $\frac{GMm}{R^2}$. The force is proportional to $1/r^2$. With an exponent of 2 in the denominator here, a distance that is twice as far away yields a force of gravity that is only 1/4 as strong.
    - We could imagine a universe in which this exponent was a different value—say, 1 or 3. According to Newton, a different law of force would result in different kinds of orbits.
    - For most laws of force that we can imagine, the orbits would not be closed curves in space but complicated spiral patterns.

The elliptical orbits we see in nature are a kind of signature of the law of gravity.

- The orbital speed of a circular orbit (v_{circ}) of radius r around a body of mass M is equal to $\sqrt{\frac{GM}{r}}$.
  - Notice that the circular orbit speed does not depend on the mass of the satellite (m). That is yet another expression of the principle of equivalence. Any object will move in the same way under the force of gravity.

  - The circular orbit speed does depend on the mass of the body that is being orbited (M). A more massive planet with stronger gravity requires a faster orbit.

  - A circular orbit speed also depends on the distance. A wider circular orbit requires a slower speed. The formula here is almost identical to the formula for escape speed. At a given distance, the escape speed is $\sqrt{2}v_{circ}$. Roughly speaking, the circular orbit speed is about 70% of the escape speed.

- We can use the equation for circular orbit speed to calculate some interesting things. For example, the Earth follows a nearly circular orbit around the Sun. The Sun's mass is about 2×10^{30} kg, and the Earth's orbit has a radius of 1 AU, which is 150 million km, or 1.5×10^{11} m. Remembering the value of G, we find that the Earth's orbital speed is 29,800 m/s, or about 30 km/s.

- An orbit is a mathematical path through space. If the velocity of an orbiting object changes a little, then its orbital path also changes.
  - For example, consider a spaceship at a distance r from the center of the Earth. It is initially moving horizontally, that is, at right angles to a line from the spaceship to the Earth. If it's moving at a speed, v, that is equal to v_{circ}, then the spaceship follows a circular orbit of that radius. If v is less than v_{circ}, the spaceship follows an elliptical orbit, one that lies inside the circular path.

- The orbit still follows Kepler's first law, and the initial point in the orbit is the point furthest from the Earth on the orbit, called the apogee. On the other side of the orbit is the point closest to the Earth, called the perigee.

- An even slower initial velocity will make the perigee that much closer to the Earth. If the initial speed is too low, of course, then the orbit will intersect the surface of the Earth—the spaceship crashes.

- If the initial velocity is greater than the circular orbit speed, then the spaceship also follows an elliptical orbit, but this ellipse is larger, outside of the circular path. The starting point is now the perigee of the orbit. If the speed is great enough—greater than the escape speed—the orbit is not a closed ellipse at all. The spaceship follows a curving path into space.

Measuring an Orbit

- To describe an orbit precisely, we need to specify some key measurements. First, the major axis—the total length of the ellipse—is the sum of the pericenter distance and the apocenter distance. The semimajor axis is 1/2 multiplied by the major axis; for a circular orbit, the semimajor axis is the radius of the circle of the orbit.

- The second key measurement is e, the eccentricity, or degree of flattening, of the ellipse. A value of $e = 0$ means that the ellipse is a circle; an e near 1 means that the ellipse is long and thin. Planetary orbits are usually fairly circular.

- For our purposes, the semimajor axis and eccentricity are the most important orbital elements.

A Gravitational Paradox

- If only gravity acts—if only Newton's $1/r^2$ gravitational force is important—then the orbital elements will stay constant. If some other force acts, such as thrusts from a rocket motor, the orbital elements can be changed.

- For example, the International Space Station orbits about 400 km above Earth's surface, where only traces of atmosphere are present, creating a tiny drag on the satellite. How does this air friction affect the orbit?
    - The effect of friction is to decrease the total energy of the satellite. Some of the mechanical energy of the satellite is converted into heat. The thin air and the satellite both heat up a little bit. Thus, the satellite has to shift to an orbit with a lower mechanical energy (lower KE + PE).

    - This would be an orbit with a smaller radius; the satellite slowly loses altitude. But a lower orbit is a faster orbit, so the satellite moves more rapidly in a lower orbit. This seems paradoxical: The effect of friction actually speeds up the satellite.

    - Friction reduces the satellite's energy by a small amount, say, 1000 J. That amount of the satellite's mechanical energy is turned into heat by friction. The orbit gets lower—the satellite gets closer to Earth—so its PE gets lower. In fact, the PE decreases by 2000 J. The satellite then speeds up, accelerated by the gravitational pull of the Earth, and its KE increases by 1000 J. PE decreases by 2000 J, KE increases by 1000 J, and 1000 J is turned into heat. Energy is conserved.

Halley's Comet is named after Edmund Halley, who used Newton's laws to work out the orbit of the comet when it appeared in 1682 and predicted its return 76 years later.

- If the satellite loses a little energy by frictional heating, in the long run, it goes faster. The reason for this counterintuitive behavior is found in the way gravity shapes the motion of an orbiting satellite.

Orbital Shape

- A spaceship moving at the escape speed or faster will not follow a closed ellipse around the Earth but will fly off into space. What is the shape of its orbit? Newton showed that the possible shapes of orbits are circles, ellipses, parabolas, and hyperbolas.

- Mathematically speaking, these orbits are related. A circle is a perfectly symmetric ellipse; a parabola is similar to an ellipse in which one end has been stretched to an infinite distance; and so on. The following table shows the orbital elements of concern to us for these orbits.

Shape and Type of Orbit	Eccentricity	Orbit Speed	Energy of the Orbiting Body
Circular (closed)	$e = 0$	$v_{circ} = \sqrt{\frac{GM}{r}}$	KE + PE < 0
Elliptical (closed)	$e = 0$ to 1	$v < v_{escape}$	KE + PE < 0
Parabolic (escape)	$e = 1$	$v_{escape} = \sqrt{\frac{2GM}{r}}$	KE + PE = 0
Hyperbolic (escape)	$e > 1$	$v > v_{escape}$	KE + PE > 0

- Communication satellites are launched into geosynchronous orbits, that is, an orbit that rotates once in the same amount of time Earth rotates once—24 hours. When such a satellite is first launched, it is often placed in a low parking orbit, a few hundred km up, but it must reach a circular orbit of about 42,000 km in radius. The

change from the initial orbit is accomplished through a navigation maneuver called the Hohmann transfer orbit.

- Another basic orbital maneuver is performed to allow a spaceship to fall into orbit around a distant planet, such as Mars. The spaceship initially follows a hyperbolic orbit; to enter a closed orbit, its energy must be reduced. Its rocket motors are fired straight ahead to slow the ship down; once its speed decreases to below the escape speed, it will follow an elliptical or circular orbit around Mars.

- For every kind of orbit that gravity allows, some form of Kepler's law of equal areas holds true. If you're orbiting the Earth in a spaceship, a line from the Earth's center to the spaceship sweeps out equal areas in equal times. If you're moving on a parabolic orbit around the Earth, the same thing is true. When the spaceship is close, it moves faster than when it is far away. Why?
    - Kepler's second law is really a familiar law of mechanics in disguise: the conservation of angular momentum. The angular momentum for something moving in a circle is the product of mass, speed, and distance from the center.

    - If something follows a free orbit, the angular momentum, ℓ, stays constant. If the radial distance, r, decreases, the speed, v, must increase. The product of those two, $r \times v$, determines the rate that area is swept out. Conservation of angular momentum means that the line sweeps out equal areas in equal times.

- Newton perfected Kepler's harmonic law by showing its connection to gravity. For a satellite orbiting a body of mass M, following an orbit of radius (or semimajor axis) a, with orbital period P: $a^3 = \left(\frac{G}{4\pi^2}\right)MP^2$. This law can be used to determine the mass of objects in space, including the mass of an object in the center of our galaxy, around which stars are moving very rapidly. We cannot see this massive object directly, but using the harmonic law, we know that it is a compact body about 3 million times more massive than the Sun—a black hole.

Suggested Reading

Doody, *The Basics of Space Flight*, chapters 3–5.

Gamow, *Gravity*, chapter 4.

Questions to Consider

1. Consider Newton's cannon thought experiment. If the cannonball is fired fast enough, it goes in a circular orbit around the Earth (ignoring air resistance). What happens if the cannonball is fired a little faster than this? What happens if it is fired twice as fast?

2. Here are three types of spacecraft missions from Earth to Mars: To fly swiftly past Mars, to enter into orbit around Mars, and to land gently on the surface of Mars. Which one requires the most fuel? Which one requires the least? Explain why.

Problem

In this lecture, we saw that the speed of a body in a circular orbit of radius r around a planet or star of mass M is $v_{circ} = \sqrt{\frac{GM}{r}}$. In each orbit, the satellite traverses a total distance of $d = 2\pi r$. A body moving with a speed v travels a distance d in a time $t = \frac{d}{v}$. From these pieces of information, derive the harmonic law for circular orbits.

What Are Tides? Earth and Beyond

Lecture 8

In the last lecture, we discussed orbital mechanics—how Newtonian gravitation shapes the orbital motion of two bodies, such as a planet orbiting the Sun or a satellite orbiting a planet. We saw that the simple orbital paths are conic sections: circles, ellipses, parabolas, hyperbolas. The shape of the orbit and the motion along the orbit are determined by the energy and angular momentum of the orbiting body. In this lecture, we'll see how gravity explains a quite different phenomenon, one that's mysterious but also familiar: the tides.

The Tidal Effect

- Newton's *Principia* contains the correct explanation for the tides—a brilliant application of the idea of universal gravitation.

- Consider the fall of a bushel of apples—a spherical cloud of apples. The apples fall almost at the same rate.
  - Remember that all the apples are accelerating toward the center of the Earth; thus, the apples on the right and the left are accelerating in slightly different directions. They draw together very slightly. The apples on the top and the bottom are at slightly different distances from the center of the Earth; thus, they accelerate at slightly different rates and draw apart.

  - The net effect, which is small, is that the cloud of apples changes shape as it falls.

  - The gravitational force is slightly different in different places. This fact shows up as a mysterious force that squeezes in from the side of the cloud of apples and stretches out along the up-and-down direction.

- The tidal effect is that relative accelerations occur due to small differences in gravity. If gravity were perfectly uniform, there would be no tidal effect.

- Recall that at a radius r from a body of mass M, there is an acceleration of gravity, g, that equals $\frac{GM}{r^2}$. Near the Earth's surface, g is about 10 m/s^2; this varies in magnitude and direction depending on location.

- Let's consider two apples that are near each other, one falling at a distance h above the other. With the tidal effect, the apples are drawing apart as they fall. The relative acceleration, a, is equal to $\frac{GM}{r^2}\left(\frac{2h}{r}\right)$. There's a sideways squeezing, but it's half as big. This relative acceleration is generally much smaller than the acceleration of gravity. For two objects near Earth that are separated by 1 m, the tidal acceleration is less than 1 millionth of g.

- Certain effects would increase tidal acceleration, including a mass (M) larger than the Earth, a smaller radius (because this effect is proportional to $1/r^3$), and a larger h (separation between the two bodies).

The Rise and Fall of the Oceans

- How does the tidal effect explain the rise and fall of the oceans? As we saw, the Moon, in its orbit, is falling toward Earth. Its path is continually bending toward Earth. That acceleration is due to gravity, but the Earth is also falling toward the Moon as a result of the gravitational pull of the Moon. The Moon attracts all parts of the Earth, but different parts are attracted differently.

- Imagine a planet on which there is a uniform ocean with no land masses. As the planet and its moon orbit each other, the planet accelerates slightly toward the moon.
  - Different parts of the planet are pulled a little differently. Water is free to move around on the surface, so it responds more strongly. The ocean, therefore, is slightly deformed. It stretches

© Ingram Publishing/Thinkstock.

The essential physics of tides is that the gravitational force is slightly different in different places.

out along the line between the planet and the moon. There are two tidal bulges: one toward the moon and one away from the moon.

- The near bulge occurs because the ocean is pulled more strongly there than at the center of the planet. The far bulge occurs because the ocean is pulled less strongly than at the center of the planet.

- If the planet is rotating, it will pass through two bulges each day—or almost two because the moon is also moving in its orbit, which increases the time between successive high tides.

- According to Newton, tides are the result of differences in the Moon's gravitational pull on different parts of the ocean. Of course, there are some complications to this picture, such as tides in the atmosphere, the presence of land masses, and tides due to the Sun.

Other Tidal Effects

- Tidal effects also influence the Earth's rotation. The Earth is not a perfect sphere; it's oblate. It's wider at the equator than it is through the poles. The tidal effect of the Sun and Moon is to cause the axis of the Earth's rotation to precess like the axis of a top. That precession takes about 26,000 years for the axis to go around once.

- In addition, the gravitational tidal effect gradually slows the Earth's rotation. As the Earth rotates, water flows around to make the tides, but there is friction in the flow, and that friction leads to a loss of energy. The tidal bulges produced by the Moon lag behind the tidal forces. They can't quite keep up with Earth's rotation.
  - In a simplified picture, the tidal bulges in the ocean are at an angle to the line between the Earth and the Moon that is about 30°. That skewed tidal bulge produces small off-center gravitational forces.

  - The gravitational forces experience action and reaction. They slow the Earth's rotation, and they transfer energy and angular momentum to the Moon's orbit so that the Moon gradually recedes from the Earth.

- If you study the Earth's rotation with atomic clocks, you'll find that the Earth day is getting longer by about 1.5 milliseconds (ms) per century. Laser reflectors left on the surface of the Moon by the *Apollo* astronauts show that the Moon is receding from Earth by about 3.8 centimeters (cm) per year.

- Tidal sediments in the fossil record show us that in the Devonian period (400 million years ago), there were about 400 days per year, about 22 hours per day, and about 28 days per lunar month. The Earth was rotating faster, and the Moon was closer. About 1 billion years ago, the Earth day was about 17 hours long, and the Moon was at only about 90% of its current distance.

- The Earth-Moon system is about 4.5 billion years old. We now believe that about 4.5 billion years ago, the Moon was formed by a collision between the young Earth and another protoplanet. This collision completely destroyed the protoplanet and knocked material from the Earth out into space, which coalesced into the Moon.
  - After the collision, the Earth was rotating once every 6 hours, and the Moon was only about 25,000 km away.

  - At that distance, the tides on Earth were more than 1000 times greater than they are today. The process of tidal friction that slowed the Earth and moved the Moon's orbit was initially very rapid.

- Throughout history, this slowdown and recession rate has varied quite a bit as a result of continental drift. The placement of the continents varies the amount of friction in the ocean tides. With the arrangements of the continents as they are today, there's fairly strong friction and the slowdown rates are fairly high, but in other eras, there has been less friction and the slowdown rates have been lower.

- Tides on the Earth due to the Moon have had a great effect over time. What about tides on the Moon due to the Earth? The Moon has no ocean or atmosphere, but body tides should be much stronger. After all, the Earth's mass is 80 times the Moon's mass.
  - Tidal forces plus internal friction have slowed the Moon's rotation such that the Moon is now locked. Its rotation period is equal to its orbital period. This means that the tidal bulges stay in the same place; the situation is stable because there are no more frictional effects.

  - This tidal locking is common in the solar system. Moons of other planets, such as Mars, Jupiter, Saturn, and beyond, mostly keep one face to their planets all the time.

- In the far future, the rotation of the Earth will slow and the Moon will recede until both the Earth and the Moon are locked, always facing each other as they go around.

The Roche Limit

- Because gravity varies from place to place, a moon orbiting a planet experiences a slight distortion. It's stretched in one direction and it's squeezed in others. The planet also experiences a tidal effect due to the moon. Combined with a little friction, the tidal effect affects orbital and rotational motion.

- Tidal effects are much stronger closer to the planet. The relationship is $1/r^3$; two times closer means that gravity is four times stronger, but the tidal effect is eight times stronger. At a close enough distance—the Roche limit—tidal forces can even pull the moon apart.
  - In the simplest case, where the density of the planet and the satellite are about the same, the Roche limit is about 2.4 times the radius of the planet.

  - If you have a satellite that is held together by its own gravity and comes closer than 2.4 times the planet's radius, it will be pulled apart by tidal effects.

- The most beautiful illustration of the Roche limit is Saturn's rings, which are made of millions of small pieces of ice, all pursuing independent, nearly circular orbits around Saturn. If we superimpose Saturn's Roche limit on an image of the rings, we notice that the main rings all lie within this limit.

- The most dramatic example of tidal effects in our solar system is seen in the moon of Jupiter called Io. This moon is about the same size and mass as our Moon and about the same distance from Jupiter as our Moon, but its orbital speed is much faster because Jupiter's gravity is much greater.
  - Because of Jupiter's enormous mass, tidal effects are much stronger for Io. They're as strong for Io as the tidal effects on

Earth in the earliest days of the Earth-Moon system. But the tidal stretch is not enough to pull Io apart.

- Io's orbit is almost but not quite circular. The moon's eccentricity is about 0.004. That means that as Io orbits Jupiter, its distance from the planet varies a little bit less than 1%, which in turn means that the strength of the tidal effect varies. Io is stretched and relaxed repeatedly every couple of days.

- This repeated stretching generates heat in the moon's interior and causes Io to be the most volcanically active body in the solar system. At any given time, Io is experiencing several gigantic eruptions—driven by gravity.

Suggested Reading

Beatty et al., *The New Solar System*, chapter 16. Several other chapters, such as those on Jupiter's moons, contain striking illustrations of tidal effects.

Galileo, *Dialogue Concerning the Two Chief World Systems*, "Fourth Day." Galileo's (incorrect) theory of the tides is vividly explained in the "Fourth Day."

Newton, *Philosophiae naturalis principia mathematica*, Propositions XXXVI and XXXVII, in *On the Shoulders of Giants* (Hawking, ed.). Newton's (correct) theory of the tides is more obscurely explained in Propositions XXXVI and XXXVII of the *Principia*.

Questions to Consider

1. Planet A has a single moon. Planet B has a moon that is twice as massive but orbits twice as far away. Which planet has greater tides?

2. By Newton's third law, the gravitational pull of the Earth on the Moon is equal to the gravitational pull of the Moon on the Earth (though they are opposite in direction). Is the same equality true for the tidal effect of the Moon on the Earth and the Earth on the Moon?

Problem

In the science fiction story “Neutron Star” by Larry Niven, the hero, Beowulf Shaeffer, flies his spaceship just 2000 km (2×10^6 m) from a neutron star, a very compact body about the mass of the Sun (2×10^{30} kg). At the risk of spoiling the story, Shaeffer has a lot of trouble with the tidal effect! Calculate the relative tidal acceleration of two apples placed at Shaeffer’s head and feet, which are 2 m apart.

Nudge—Perturbations of Orbits

Lecture 9

The Moon's orbital path shifts by more than 1° each orbit. If Earth and the Moon were the only bodies in the universe, this could not happen. But other bodies in the universe, particularly the Sun, exert gravitational forces on the Earth and Moon, and those forces change the Moon's orbit. This description leads us to what physicists call the three-body problem: Three bodies move in space, exerting gravitational forces on one another. How do they move? This problem is incredibly complicated and has no neat solution. In the next three lectures, we'll look at the effects of gravity on the motions of more than two bodies, beginning in this lecture with small effects, such as the nudge that one planet exerts on another.

Discovery of Neptune

- Up until 1781, our solar system was thought to contain the Sun, six planets (Mercury, Venus, Earth, Mars, Jupiter, and Saturn), moons, and comets. In 1781, William Herschel, an English astronomer, discovered a new planet beyond Saturn, which was eventually named Uranus.

- Over the next few decades, Uranus was observed very carefully. In 1821, Alexis Bouvard, a French astronomer, noticed that Uranus does not seem to follow its orbit exactly in the way Newton would predict. In a simplified view, Uranus seems to move too fast for part of its orbit and too slow for another part.

- Two possible explanations of this phenomenon arose: (1) Perhaps Newton's law of gravitation is only approximate at great distances from the Sun, or (2) perhaps there is an undiscovered planet beyond Uranus, nudging Uranus in its orbit.

- In 1845, the English mathematician John Couch Adams used an iterative process to try to figure out where the undiscovered planet must be. He started out with an assumed orbit for the unknown

planet, then calculated its effect on Uranus using Newton's laws. He then adjusted the unknown orbit to improve the agreement between the calculations and the known data on Uranus.

- In 1846, the French astronomer Urbain Le Verrier independently took up Bouvard's work. He followed the same path as Adams, but he arrived at a much more definite solution. He communicated his ideas to Johann Galle, an assistant at the Berlin observatory, who almost immediately found a new planet—Neptune—based on Le Verrier's prediction of its location.

- The discovery of Neptune was a tremendous vindication of Newton's law of universal gravitation. Tiny variations in the orbit of Uranus meant that either Newton was wrong or there was an unknown planet; in fact, the planet was found. This discovery was the first time that gravity had been used as a tool to discover a new planet. Adams and Le Verrier were able to deduce the existence of something new in the universe just by seeing the effects of its gravity on a known object.

- Gravity is the perfect tool for letting us see the unseen because everything in the universe exerts gravity and gravity affects everything in the universe. Gravity has been the idea behind more than one amazing advance in our knowledge of the universe.

The Doppler Effect

- Planets orbiting other stars beyond our Sun are called exoplanets. Consider an exoplanet orbiting a massive star.
    - Actually, both the planet and star orbit each other, both orbiting a common center point. Because the star is more massive than the planet, that center point is close to, or perhaps even inside, the star.

    - The star wobbles around as the planet follows its orbit; what produces that wobble is the gravitational force of the planet on the star.

- This force is equal and opposite to the gravitational force of the star on the planet, but because the star has so much more inertia, it doesn't move much. The wobble is so small and the stars are so far away that any side-to-side motion ("proper motion") is too small to see, but motion toward us and away from us ("radial motion") can be detected by careful measurements using the Doppler effect.

- When we look at the spectrum of light from a star, we find that certain specific wavelengths are darker than expected. These wavelengths are absorbed by gases in the star's outer envelope. The wavelengths are called spectral lines, and they often appear as dark vertical lines on the continuous rainbow spectrum from the star.
    - These lines are characteristic of the elements making up the star, the same elements we have in laboratories here on Earth. We know exactly what wavelengths of light they represent, but the spectral lines from a star might be shifted. The wavelength might be smaller or greater because the star might be moving toward us or away from us. This is called the Doppler effect.

    - If the star is moving toward us, the lines are shifted toward the blue end of the spectrum ("blueshift"). If the star is moving away from us, the lines are shifted toward the red end of the spectrum ("redshift").

    - If we see the lines of a star change—from redshift to blueshift repeatedly—over a period of days, weeks, or months, then we know that the star is wobbling in space. From that wobble, we can deduce the existence of an unseen body orbiting the star, nudging it this way and that. In other words, we can deduce the existence of a planet.

- The idea of using the Doppler effect to find planets around other stars is called the radial velocity method, which tells us the period, size, and shape of the orbit and the mass of the planet. Since the

mid-1990s, astronomers have used the radial velocity method, together with some other techniques, to find hundreds of planets beyond our solar system.

Orbital Change

- As we've seen, the planets orbit the Sun, but they exert gravitational forces on one another, continually nudging one another so that the orbits are slowly modified. This fact worried Newton, who thought that after a long time, the orbits of the planets would get out of sync, and the planets might even run into each other.

- A century after Newton, Pierre-Simon Laplace, a great French mathematician, concluded that the orbits in our solar system were fairly stable. Modern computer calculations confirm that the chances of a planetary collision are low.

- Why don't the orbits of the planets change more? Most of the changes we expect in the orbits are oscillatory. In other words, the orbits do not simply change in one direction without bound. The orbit of Mars, for example, might become a little more eccentric for a period of time; the ellipse might be slightly elongated as Mars is nudged around by other planets. Then, after a few million years more, the nudges will tend to make the orbit more circular again. Everything is continually changing, but those changes generally stay within bounds.

- The Earth's orbit, which seems so constant, also changes slightly over time in one way or the other, and these changes have had important effects. For example, over the last few million years, the Earth has experienced repeated glaciations—ice ages. These occur every couple of periods of 100,000 years. We now think that these repeated ice ages have been triggered by small periodic variations in the Earth's orbit and rotation due to the gravitational influence of the Moon and other planets in the solar system.

- Mercury has a fairly elliptical orbit. Its eccentricity is 0.21. The major axis of the orbit of Mercury does not stay still but gradually

turns in space. That means that the perihelion point, the point of closest approach to the Sun, precesses very slowly. In a century, it moves about 5600 seconds of arc (arcsec), which is 1.5°.

- Almost all of this movement can be accounted for by the gravitational influence of other planets in the solar system. As other planets nudge Mercury, its orbit should precess by more than 5500 arcsec per century, but this leaves 43 arcsec per century left over, 1/80 of 1°. That's the extra shift in the perihelion point of Mercury over 100 years. It's extra once we take into account the gravity of the known planets and Newton's laws and so on.

- Even though this perihelion shift is tiny, it's terribly important, because it tells us that something unknown is going on.
    - In the 1850s, when this problem of Mercury's orbit first came to light, Le Verrier posited that there must be an unknown planet exerting some extra influence on Mercury. He predicted that there was a small planet, given the name Vulcan, orbiting inside Mercury's orbit.

    - If this planet existed, however, it would have to cross the face of the Sun from time to time, where it could be spotted by observers on Earth.

    - Le Verrier spent the last 20 years of his life trying to find evidence of a transit of Vulcan, but his search was to no avail. Even after decades of searching, astronomers had to admit that no such planet actually exists.

- That leaves us with a mystery: the tiny, uneradicable mystery of the perihelion of Mercury, which moves around 1% faster than it should. That extra shift is only 1° every 8000 years. Nevertheless, from the point of view of Newtonian physics, it is an anomaly. The mystery of Mercury was eventually solved by no less than Albert Einstein, but the final solution to the problem required a whole new theory of gravity: the general theory of relativity.

Suggested Reading

Beatty et al., *The New Solar System*, chapters 16 and 23.

Questions to Consider

1. Anomalies in the orbit of Uranus were found to be due to the gravitation of a new planet, Neptune. Anomalies in the orbit of Mercury, however, were not due to a new planet. In one case, the puzzle could be resolved within the existing (Newtonian) theory of gravity, but in the other case, it could not. This brings up one of the most important questions in science, which you should consider: How can we tell whether a piece of anomalous data requires us to change our basic theory?

2. The first exoplanets discovered by the radial velocity method were so-called "hot Jupiters"—very massive planets extremely close to their parent stars. However, such planets are believed to be unusual. If they are actually so rare, why were they discovered first?

Resonance—Surprises in the Intricate Dance

Lecture 10

A computer study of the long-term stability of planetary orbits in our solar system has shown that those orbits will likely be stable for billions of years. There is a small chance, however, that something catastrophic might happen: Mercury might collide with Venus or go careening off in the solar system; other planets might be affected. The main culprit behind this possible destabilization of the solar system is Jupiter, specifically, the possibility that Mercury and Jupiter might fall into an orbital resonance. In this lecture, we'll talk more about the three-body problem but not just its tiny effects. We'll see striking patterns and large-scale results that emerge from three-body interactions.

Defining Resonance

- Resonance is said to occur whenever a small periodic force produces a large effect on a periodic motion; resonance is a general phenomenon, not just an effect of gravity and orbital mechanics.

- The first ingredient for resonance to occur is periodic repeated motion, such as the movement of a child on a swing. Left to itself and neglecting friction, that swinging motion would repeat over and over. In the real world, the natural period of swing is about 4 s.

- The second ingredient for resonance is a small repeated periodic force. In our example, we periodically give the child a push. Each push is the same, in the same direction, and the pushes have their own timing. Suppose we push with the natural swing period, every 4 s. Then, each push happens at the same point in the swing, and over time, the effect of the pushes adds up; that's resonance.

- Suppose we push every 3 s or every 5 s. Sometimes we're pushing the child faster, but sometimes we're slowing the child down; over time, the effects of such pushes tend to cancel out; the pushing is out of resonance.

The two ingredients for resonance are periodic repeated motion and periodic repeated force—both present in the action of pushing a child on a swing.

- Suppose we push every 8 s, that is, every other swing of the child. Each push happens at the same point in the swing, every other swing. Because the pushes have a cumulative effect, this is still a kind of resonance, but in this case, it would be a 2:1 resonance.

- Resonance can happen when a periodic motion is affected by a periodic force if there is a simple ratio between the two periods, a 1:1 or a 2:1 ratio or something a little more complicated. In such situations, a small force can have a very large cumulative effect.

- The phenomenon of resonance is everywhere; it's used, for example, in the oscillating electric circuit that allows a radio receiver to tune into a transmitter hundreds of miles away.

Resonance with Jupiter

- Let's return to orbits and gravity. Consider the Sun and the planet Jupiter. Jupiter is more massive than all the other planets combined,

but it's still only 1/1000 the mass of the Sun. Orbiting the Sun are thousands of asteroids that are much smaller than any planet. Many asteroids orbit in the asteroid belt, a region of the solar system between the orbits of Mars and Jupiter.

- Jupiter's gravity tugs on asteroids in this belt. The strongest pull happens when the planet is closest, that is, when both the planet and an asteroid are on the same side of the Sun. From our discussion in the last lecture, we might think that the long-term effect of this pull would be small. The small nudges generally cancel out in the long run. But there might be a resonance at work. We have periodic motion—the orbit of the asteroid—affected by a periodic force—the pull of Jupiter. The real question is: What is the timing?

- The orbital period of the asteroid is less than the period of Jupiter's orbit because its orbit is smaller than Jupiter's. Suppose that the asteroid has just half of Jupiter's orbital period. That happens for an asteroid whose orbit is around 3.3 AU. This means that the asteroid is in a 2:1 resonance with Jupiter.
    - When Jupiter and the asteroid are aligned on the same side of the Sun, the asteroid experiences a small tug of gravity. The gravity of Jupiter is continuous, of course, but it is strongest when the planet is closest. After one full orbit, the asteroid is back to the same point in space. Jupiter is on the other side of the Sun, far away.

    - After another orbit, the asteroid is again back at the same point, but now it is aligned with Jupiter. The tug of gravity from Jupiter exactly repeats itself at exactly the same place in the asteroid's orbit.

    - Over time, the effect accumulates. The orbit of the asteroid changes significantly. It's difficult to predict what will happen after a long time. The asteroid might run into a planet or another asteroid, or it might end up in an orbit that is no longer in resonance with Jupiter. Whatever happens, an asteroid in a 2:1 resonance with Jupiter will not stay in that orbit for long.

- On a graph of the asteroid belt population, the horizontal axis shows the radius of an asteroid's orbit in AU. Kepler's harmonic law tells us that the radius of the orbit is related to the orbital period. The vertical axis shows the number of asteroids with orbits of any given size. The resulting graph has some very definite features.
    - The outer edge of the main asteroid belt lies at about 3.3 AU, exactly at the 2:1 resonance point with Jupiter. Almost no asteroids have an orbit there, and there are relatively few asteroids further out.

    - Another dip in the graph occurs at about 2.5 AU, near the 3:1 resonance point with Jupiter. Many asteroids have orbits a little larger or smaller than this, but there are almost none right at that radius.

    - Other obvious dips in the population graph occur at the 5:2 resonance point and the 7:3 resonance point. These are, in effect, gaps in the asteroid belt; these so-called Kirkwood gaps are due to orbital resonances with Jupiter.

Orbital Resonance in Action

- In some cases, orbital resonance can make certain orbits especially stable. The members of the Hilda family of asteroids, for example, are not very close together, but they all have orbits with a semimajor axis of about 4.0 AU, putting them in an exact 3:2 resonance with Jupiter. Because of timing, this resonance actually protects the Hilda asteroids from strong perturbations by the gravity of Jupiter. When a Hilda asteroid is at its farthest point from the Sun, closest to Jupiter's orbit, the planet itself is always far away from the asteroid.

- The rings of Saturn represent an even more spectacular example of orbital resonance. The gap between ring A and ring B is known as the Cassini division, and it's caused by a 2:1 orbital resonance of bodies at that point with a moon of Saturn called Mimas.

- The rings of Saturn are also a perfect place to see other delicate effects of gravity. One example is the F ring, a narrow ring just outside the A ring. The gravity of two small moons, Pandora on the outside of the ring and Prometheus on the inside, exerts opposing forces on bodies in the F ring. The net effect of the tiny gravitational forces of these small shepherd moons is to hold the particles in the F ring together.

- Two other moons of Saturn, Janus and Epimetheus, share almost the same orbit, but the small gravitational pull they exert on each other causes them to continually switch places, shifting between a slightly larger outside orbit and a slightly smaller inside orbit. This pattern is known as a horseshoe orbit.

Lagrange Points

- Lagrange points serve as an important example of the idea that in addition to perturbing the orbit of a small body, gravity can also keep that orbit stable for a long time.

- Consider an orbiting system of two large bodies, the Earth and the Moon. A third body, a satellite, will change its location, of course, relative to the Earth and Moon as it orbits, but Lagrange points are points where a satellite can stay fixed in this system relative to the Earth and Moon. These points are designated L1 through L5. Of course, the satellite is not really stationary but, in effect, is moving in formation with the Earth-Moon system.
    - For example, L1 is on the line between the Earth and Moon, but it's closer to the Moon. Ordinarily, a satellite orbiting the Earth closer than the Moon would orbit faster, but the Moon's gravitational pull effectively cancels some of the Earth's pull. The satellite actually orbits with the same period as the Moon.

    - L2 is also along the Earth-Moon line but outside the Moon's orbit. Ordinarily, this orbit would be slower, but the Moon's gravitational pull effectively adds to the Earth's pull so that

the satellite orbits with the same period as the Moon. It stays behind the Moon as seen from Earth.

- L3 is also along the Earth-Moon line but on the other side of the Earth. It's just inside the Moon's orbit.

- L4 and L5 are on the Moon's orbit, but they are 60° ahead of and behind the Moon. That is, the Earth, Moon, and satellite form an equilateral triangle.

- All five Lagrange points orbit the Earth synchronously with the Moon.

- Suppose a satellite is close to but not quite on a Lagrange point. It's moving nearly but not quite at the right velocity. What would happen?
  - The L1 through L3 points are unstable. If the satellite is not exactly in the right place with the right velocity, it will drift farther away.

  - L4 and L5 are stable. If the satellite starts close to one of these points, it may drift around in a complicated way, but it will stay close.

- L1 through L3, though unstable, can still be useful. In the early 2000s, a space probe called the Wilkinson Microwave Anisotropy Probe (WMAP) was positioned close to the L2 point of the Earth-Sun system. This allowed it to orbit the Sun more than 1 million km from Earth but also to stay close to Earth to enable communication. Because L2 is unstable, the spacecraft had a tendency to drift away over a period of weeks; thus, it used small rocket thrusters to keep it in place.

- It's no surprise to find this phenomenon at work in Saturn's system. There are small moons at the L4 and L5 points of many of Saturn's large moons. We also see many asteroids at the L4 and L5 points of the Sun-Jupiter system.

- The stationary points that Lagrange discovered are not just mathematical curiosities. They are real features of our own solar system. Doubtless, they are real features of other planetary systems, as well.

Suggested Reading

Baez, "Lagrange Points."

Tyson, *Death by Black Hole*, chapter 9.

Questions to Consider

1. Think of several ways that Jupiter's gravity affects asteroids in the solar system.

2. Imagine a distant binary star system, where a smaller star orbits a larger star. At which of the Lagrange points in this system would you look for planets and asteroids?

3. The asteroid 3753 Cruithne orbits the Sun in a horseshoe orbit relationship with the Earth. What is the average orbital period of 3753 Cruithne around the Sun? Some have described this asteroid as "Earth's second moon." Do you agree with this description?

The Million-Body Problem

Lecture 11

As we've seen, the three-body problem is incredibly complicated, but what can we say about more bodies interacting via gravitation? Imagine hundreds of stars in a cluster or billions of stars in a galaxy. At first, these situations seem to be far too complicated to be understandable. But surprisingly, we can say useful things even amid such complexity. Important general facts hold for 3 bodies or for 3 trillion bodies. Understanding those facts—those subtle consequences of the law of universal gravitation—will help us understand the formation of everything from planets to galaxies and will once again tell us that there is more to the universe than we can see.

Gravitation in Star Clusters

- Three mathematical theorems of Newtonian mechanics can help us understand the expansion or contraction of a cluster of stars, all moving in complicated orbits. Two of these principles are familiar to us, but the third one is new.

- The first principle is the conservation of angular momentum. If the cluster has a net rotation about some axis, that rotation persists. If the cluster contracts, the rotation speeds up. If the cluster expands, the rotation slows down.

- The second principle is the conservation of mechanical energy. The total mechanical energy, KE + PE, stays constant. The total KE is simply the sum of the KE for each star. Faster motion means a higher KE. Slower motion means lower KE.

- PE is a bit more complicated. It's associated with gravitational forces.
    - o If we have two stars at a distance r apart, then PE = $-\dfrac{Gm_1m_2}{r}$. Remember, gravitational PE is negative, and it tends toward zero if the two masses are very far apart.

- If we have thousands of stars, the total PE is a sum of similar terms, one term for each pair of stars, and that sum can be quite large. For 100 stars, there are almost 5000 pairs of stars; for 1 million stars, there are about 0.5 trillion pairs of stars.
- The PE of a cluster stars with a total mass M and radius R is approximately equal to $-\frac{GM^2}{R}$. This estimate is simple, but it ignores how the stars are distributed within the cluster.

- For a dense, compact cluster of stars, R is small; thus, PE is lower—more negative. For a sparse, widespread cluster, R is large; thus, PE is higher—less negative. If the stars are widely scattered through space, PE is close to zero.

- The principle of conservation of energy tells us that for an isolated cluster of stars, KE + PE stays constant. If the cluster shrinks, PE becomes more negative (decreases), and KE must increase. The stars must speed up. If the cluster expands, PE becomes less negative (increases), and KE must decrease. The stars slow down.

- If the total energy is greater than zero, then the cluster will disperse into space. If the total energy is less than zero, then the cluster is held together by gravity.

The Virial Theorem

- The third principle at work here is the virial theorem, a theorem about the average behavior of a system depending on what forces are acting. We'll use a version of this theorem that applies to gravitational forces.

- Our system consists of N bodies moving under mutual gravitational attraction. It's gravitationally bound together, so KE + PE (total energy) is less than zero. KE is positive, but PE is negative.

- For the virial theorem, we consider a new combination of energies. We take 2KE + PE; that is, we now count KE twice as much. This quantity is not conserved, but it has a remarkable

property. According to the virial theorem, over time, this sum, on average, is zero.

- In other words, the average speed of motion given by KE is related to how densely clumped together the cluster is, given by PE.

- If speeds are too fast, then 2KE + PE is greater than zero. The system will tend to expand. If speeds are too slow, then 2KE + PE is less than zero, and the system tends to contract. If the speeds are just right, 2KE + PE is equal to zero, the system stays about the same over time.

- For our purposes, we will call the result of 2KE + PE the fugacity (f). According to the virial theorem, the value of f tells us whether or not a system expands. It expands if f is positive; it contracts if f is negative. In the long run, f is about zero on average.

- A computer example shows us what happens with a cluster of 1000 particles.
  - At the start, the particles are moving slowly, so f is negative. At first, the system contracts; PE becomes more negative.

  - KE increases as the stars speed up. The total energy stays the same, but f now becomes positive. The system expands.

  - The system oscillates for a time, but in the long run, most stars become part of a denser cluster. On average, f is about zero.

Finding Dark Matter

- In 1937, a Swiss astronomer, Fritz Zwicky, applied the virial theorem to the Coma cluster of galaxies, consisting of about 1000 galaxies.

- In this cluster, f is much greater than zero. KE is much too large. But it's unlikely that the Coma cluster is dispersing. Zwicky posited that only part of the mass of the cluster is visible; if that's true, then the estimates for KE and PE would be wrong, and these discrepancies

would throw off the virial theorem.

The Coma cluster consists of about 1000 galaxies; Fritz Zwicky calculated that the apparent excess of kinetic energy in the cluster indicates the presence of some mass that is not visible to us—dark matter.

- Zwicky realized that we seem to see too much KE in the Coma cluster, which means that there is mass in the Coma cluster that is not visible to us—what astronomers have come to call dark matter. In the 1970s, the American astronomer Vera Rubin confirmed Zwicky's discovery in her studies of galactic rotation in spiral galaxies.
  - o The typical orbital speed in the outskirts of a visible disk of a galaxy is about 200 km/s. From this, we can calculate that spiral galaxies often have about 100 billion times the Sun's mass.
  - o Rubin found that the orbital speed of stars beyond the edge of the visible disk of the galaxy is about the same, although we might expect that their orbital speeds would be much slower, given their distance from the galaxy center. The larger orbit, therefore, must enclose more mass.
  - o This means that the visible galaxy is only a small part of the whole. Most of the mass of the galaxy is in a large invisible halo of dark matter extending into space beyond the visible edge of the galaxy.
- Both Zwicky and Rubin raised the possibility that Newton's law of gravity is not quite correct for large systems, such as galaxies and

galaxy clusters. Physicists call this idea MOND, meaning "modified Newtonian dynamics."

- If dark matter exists, we might be able to roughly map it out by studying how gravity bends light (gravitational lensing). In contrast, if MOND is correct and we made maps of light bending, there would be an apparent halo of dark matter around every large object, such as a galaxy or a cluster, because the dark matter there would be only a shadow of regular matter. The two could never be separated.

- A computer model of a galaxy cluster under ordinary Newtonian gravity shows that dark matter comprises most of the cluster. Assuming the existence of dark matter, if this cluster were to collide with another, then the gas and the dark matter in the clusters could be separated. This separation is a key piece of evidence in favor of Newtonian gravity and the existence of dark matter.

Formation of the Universe

- Another application of the virial theorem tells us how things in the universe first formed. This application was originally worked out more than 100 years ago by the English astrophysicist James Jeans.

- Imagine we start with a cloud of gas in space. The cloud has a certain density and a certain temperature. The cloud might either disperse into empty space or experience gravitational collapse. The gravitational collapse of a tenuous cloud of gas is the first step in the formation of almost anything in the universe.

- According to Jeans, the key to this is the virial theorem, but he does not apply it to a cluster of stars or galaxies. He applies it to the individual molecules in the cloud of gas.
    - PE is still a measure of how clumped together that mass is. It's still reasonable to estimate the PE of a cloud of mass M and radius R by $-\frac{GM^2}{R}$.

- The KE of the molecules, though, is just the heat content of the gas, which is related to temperature. The KE is also proportional to the mass.

- Our original cloud will contract, expand, or stay the same, depending on whether f is negative, positive, or zero. Because gravity is very weak for a small cloud, PE is close to zero. KE dominates and f is positive; the cloud expands and tends to disperse.

- Consider, though, a cloud that is 10 times larger in radius and has the same density and temperature. Its volume is 1000 times greater, so its mass is 1000 times greater. KE is 1000 times larger, but how does PE get multiplied?
    - In our formula, PE depends on M^2/R; thus, PE is multiplied by $1000^2/10$. In other words, PE is multiplied 100,000 times.

    - For a large cloud of the same gas with the same temperature and density, the negative PE is more important; f is negative. The cloud must contract.

- Jeans worked out just how large a cloud has to be for gravitational collapse to occur. This "Jeans mass" depends on the density and temperature of the gas. If the mass of the cloud is less than the Jeans mass, the cloud will disperse. If the mass of the cloud is greater than the Jeans mass, the cloud will undergo gravitational collapse.
    - For a molecular cloud in our own galaxy with about 1000 hydrogen molecules per cm^3 and a temperature about 10° to 20° above absolute zero, the Jeans mass is several times larger than the mass of our Sun.

    - A cloud larger than that will contract, which is why the stars in our galaxy mostly form in groups.

    - Jeans's insight is that gravity determines where, how, and when tenuous clouds in deep space can contract to form stars, galaxies, and planets.

- Almost 5 billion years ago, a cloud of hydrogen and helium gas found itself on the right side of the Jeans line for gravitational collapse. Driven by its own gravitational forces, the cloud contracted to a tiny fraction of its original size and eventually formed the Sun and the planets of our solar system.

Suggested Reading

Case Western Reserve University, "Galaxy Crash JavaLab!"

Gates, *Einstein's Telescope*, chapter 1. Despite its usefulness and importance, I am aware of no generally accessible accounts of the gravitational virial theorem—at least, not until now! (This is, of course, a standard topic in upper-level astrophysics textbooks.) The dark matter discoveries of Zwicky and Rubin are well-described by Gates.

Questions to Consider

1. Why does the virial theorem not apply to a solid body, such as the Earth? Think about what forces are acting on the particles that make up the Earth.

2. A cloud of interstellar gas is just on the verge of contracting due to gravity; that is, its mass is not quite the Jeans mass. Predict what is likely to happen if the cloud gets colder, warmer, larger in radius, or smaller in radius.

Problem

The virial theorem says that PE = –2KE for a galaxy cluster. Suppose only 1/3 of the mass of a galaxy cluster is visible. If we write PE′ = –κKE′ for the visible potential and kinetic energies, what is the coefficient κ? Is it greater or less than 2? If we include only the visible matter, does the total energy, KE′ + PE′, appear to be positive or negative?

The Billion-Year Battle

Lecture 12

At the end of the last lecture, we saw how gravity determines when we can expect the gravitational collapse of a tenuous cloud of gas. That's how planets, stars, and even whole galaxies form. Of course, the Sun and the planets do not continue to contract. Other internal forces balance the inward pull of gravity. That balance is called hydrostatic equilibrium. For solid planets, the collapse is halted by interatomic forces; the Sun, though, is much larger and made of gas. There's a continual struggle between gravity pulling inward and pressure forces pushing outward. In this lecture, we'll see how those forces work in the Sun and how that struggle will turn out in the future.

Pressure Forces and Density

- Pressure is a contact force, exerted between things that are touching, such as the force between your shoes and the floor. This kind of force is not only between solid objects, but it can also be exerted between different parts of a fluid. Pressure forces are due to short-range forces between molecules.

- Pressure is measured as force per unit area; it's measured in N/m^2, which is a unit called a pascal (Pa). The pressure of air at sea level is about 100,000 Pa, which is sometimes called 1 atmosphere (atm) of pressure.

- In any condensed body held together by gravity, the internal pressure—an outward force—must balance the gravitational force, which is an inward force. This is called hydrostatic equilibrium.

- The pressure at sea level is 100,000 Pa. The downward force of the air must equal the upward force of the water at the water surface. The water supports the weight of the air above it.

- Consider the top 10 m of water in the sea. For each m^2 of the sea, there's a downward pressure force of 100,000 N at the top from the pressure of the air.

- We also have to consider the weight of the water. We're thinking about a column of water 1 m^2 by 10 m tall. That's 10 tons of water. That column pushes down with a force of gravity equal to another 100,000 N. The pressure force at the bottom of the layer, 10 m deep, must be 200,000 N/m^2—2 atm.

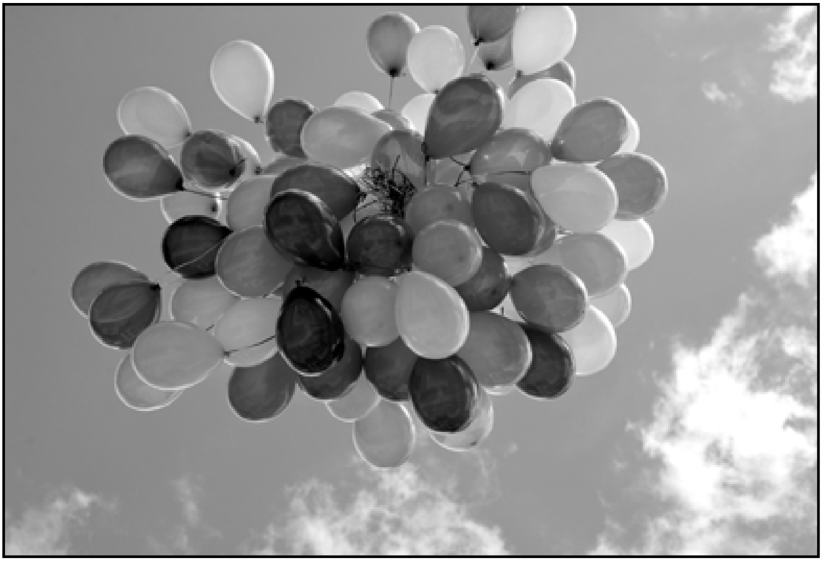
- Each layer of the ocean must support the weight of all the layers above it. The deeper the layer, the greater the pressure must be. Roughly speaking, each 10 m of depth adds 1 atm. This is a requirement of the balance between pressure and gravity.

- Water stays at nearly the same density as pressure increases. Compressed water at the bottom of the ocean is only a few percent denser than at the surface. Each 10-m layer of water adds almost the same extra weight. We have the same pressure increase for each 10 m of depth.

• Hydrostatic equilibrium also applies inside the Earth. Rock layers are many times denser than water. The pressure at the center of the Earth is more than 3 million atm. This pressure must literally hold up the weight of the Earth.

• The same reasoning applies to the Earth's atmosphere. What holds the air up? It must be that air pressure is greater lower down and less higher up. The greater pressure of the lower layers supports the weight of the upper layers. The atmosphere, too, is in hydrostatic equilibrium.
  - Unlike the density of water or rock, the density of air is quite variable. At high altitudes, the pressure is low, and the air is very thin.

  - Very roughly, the rule of thumb is that air pressure and density each go down by about 10% for each 1000 m of altitude.



A helium balloon has the same buoyant force—the same unbalanced pressure force—but less weight than the same volume of air; that means there's some buoyancy left over to push the balloon upward.

Pressure and Temperature

- There's also a connection between pressure and temperature. For liquids and solids, this relation is not very close. High or low pressure can occur at high or low temperatures. For gases, however, the connection is much closer. The pressure of a container of gas is, in fact, proportional to the temperature measured above absolute zero.

- A star, such as the Sun, is made of a kind of gas called a plasma. This is gas at such a high temperature that the electrons are released from their atoms. The gas becomes a mixture of positively charged nuclei and negatively charged electrons, all flying around independently, not bound together in neutral atoms.

- The Sun has a huge mass and, therefore, a huge gravitational pull. Hydrostatic equilibrium means that the internal pressure of the Sun

is also huge—250 billion atm. The gas pressure can be so high because the Sun has a very hot interior—about 15 million kelvin (K). Such a high core temperature yields immense pressures for hydrostatic equilibrium, maintaining the balance between pressure and gravity.

- The Sun radiates 400 trillion trillion J/s. We would expect, then, that its interior would slowly cool down, and internal pressure would decrease. Gravity would begin to win the battle, and the Sun would contract. Of course, the contraction would increase pressure and temperature at the core, so a new equilibrium would be established. Then, the Sun would cool some more and contract some more.

- At temperatures above about 8 million K, hydrogen nuclei can fuse together and form helium nuclei. In the process, the hydrogen atoms release a huge amount of energy. The rate at which this energy is produced depends on the rate of nuclear reactions in the Sun's core. That rate, in turn, depends on the density and pressure in the core. Higher temperatures or higher density leads to faster reactions and, thus, the release of more energy.

Battle in the Stars

- The internal conditions in the Sun are determined by hydrostatic equilibrium, in other words, by a balance with gravity. The same thing is true for other stars. Most stars are mostly hydrogen. Hydrogen fusion reactions in the core keep the interior hot, maintaining internal pressure and, thus, hydrostatic equilibrium. The interior conditions of the star are determined by gravity.

- Consider a star of a given chemical composition, mostly hydrogen and helium. Practically everything important about the star—its radius, temperature, density, brightness—is determined by the force of gravity, therefore, by the star's mass. This fact, that everything depends on the mass, is sometimes called the Russell-Vogt theorem, first recognized by the physicists Henry Norris Russell and Heinrich Vogt in the 1920s.

- As an example, consider the mass-luminosity relation. Luminosity is the total energy output of the star. The mass of the star determines gravitational forces, which determine interior conditions, which determine nuclear reaction rates, which determine energy output. More massive stars are actually much more luminous.

- Because of the relation between mass and luminosity, there is an effective upper limit to the mass of a star. A more massive star would have greater gravity, but the nuclear reactions in its core would generate far more energy. For a star more than about 150 times the mass of the Sun, the exact figure is still somewhat murky; the outward pressure is enough to overcome the star's gravity and blow the outer layers of the star into space.

- For an ordinary star, internal pressure depends on high central temperatures. Those, in turn, depend on the energy released by nuclear reactions. In the battle between pressure and gravity in a star, one side—the pressure side—relies on nuclear reactions. In other words, that side is using up ammunition; it's consuming hydrogen fuel in the core. A crucial question for any star is: How long will the ammunition last?
    - For our Sun, the hydrogen-burning lifetime is about 10 billion years. We're about halfway through that lifetime.

    - Low-mass stars burn so much less brightly than the Sun that their more limited fuel supply lasts much longer. More massive stars have more fuel, but they burn much brighter and, thus, have shorter hydrogen-burning lifetimes.

- What happens to a star when the ammunition runs low?
    - Most of the hydrogen in the core of the star has been converted into helium. A star like our Sun will settle for a while into a helium-burning stage, and its core will become much hotter. Helium fuses rapidly into carbon in the core of such a star. The core is surrounded by a shell of gas that still has hydrogen, which is now hot enough for hydrogen fusion to occur.

- The total energy release of the star will increase, which causes the outer layer of the star to expand. This is called the red giant stage of the star's life.

- Eventually, all usable nuclear fuel is exhausted; pressure's ammunition has run out. The result is core collapse and the release of a great deal of energy. This energy release may cause the outer layers of the star to be ejected into space in a huge explosion.

- A star like our Sun or smaller then becomes a white dwarf. Such a star might have the mass of the Sun but only the radius of the Earth. In other words, it has a huge density. Electron degeneracy pressure, a quantum force between electrons in the ultradense matter that composes the white dwarf, provides the internal pressure for hydrostatic equilibrium in the star. The white dwarf will cool off over time and become a black dwarf.

- Interestingly, there is an upper limit to the possible mass of a white dwarf star, called the Chandrasekhar limit. For a star above about 1.4 solar masses, even electron degeneracy pressure cannot balance gravity. A collapsing star that big cannot settle down as a white dwarf; it must keep collapsing. The result is a neutron star.
  - A 2-solar-mass neutron star is about 10 km in radius; that's 100 million times denser than a white dwarf star.

  - Newly formed neutron stars generally rotate up to 1000 times/s. Angular momentum is conserved during the star's collapse. Thus, as it collapses, its rotation speed is tremendously amplified.

- What happens to a collapsing stellar core even more massive than that of a neutron star? Gravity is too strong for electron degeneracy pressure to stop it, so there's no stopping at the neutron star stage. For the most massive stars in the universe, the battle between pressure and gravity has a winner: gravity. The final state is a black hole. That's an object so massive and compact

that not even light can escape from it. It's a place in the universe where gravity is at its most extreme.

- The ultimate fate of a star depends on its mass. That fate is set by the final standoff between gravity and internal pressure. In a white dwarf or a neutron star, gravity is balanced by quantum degeneracy forces. In a black hole, we see the final victory of gravity in the billion-year battle. At this point, our Newtonian understanding of gravity is simply not enough; we must turn to Albert Einstein for a better theory.

Suggested Reading

Thorne, *Black Holes and Time Warps*, chapters 4–6.

Questions to Consider

1. A rowboat floats in the water. If a person climbs into the rowboat, then the boat sinks a little farther into the water. Explain this in terms of the water pressure on the bottom of the boat.

2. A cork floats in water because of the buoyant force. Would there be any buoyant force if the water and the cork were in an environment without gravity?

3. Imagine a chain hanging vertically under its own weight. There is a tension force acting between the links, which holds the chain up. Is this force greater at the top of the chain or at the bottom? Compare this situation to hydrostatic equilibrium.

From Forces to Fields

Lecture 13

So far, we've seen phenomena both on Earth and in space that are governed by gravity. We've found that gravity lets us see things in the universe that would otherwise be invisible. Our turning point now is a revolutionary understanding of space, time, motion, and gravitation developed by Albert Einstein. These ideas will be our main business for the rest of the course. Occasionally, it may seem that we've wandered away from the topic of gravity. We'll talk about the speed of light, clocks and time, and so on. In these discussions, we're building a foundation. In the lectures to come, we will need every part of that foundation to build our new understanding of gravity.

Electric and Magnetic Fields

- Einstein's work grew out of a remarkable shift in physics during the 19th century—a shift from thinking about forces between particles to thinking about fields in space. This shift introduced new ideas that extended physics beyond Newton's mechanics and introduced some profound puzzles.

- Consider an example of gravity in action: a planet orbiting the Sun. We've said that the planet is steered by a gravitational force exerted by the Sun, but the Newtonian picture is more complicated than that. Everything in the universe exerts some gravitational force on the planet; thus, the planet is steered instantaneously from everywhere in the universe. This idea is known as action at a distance.

- In the early 19th century, an alternative point of view emerged in connection with electric and magnetic forces. These are long-range interactions, like gravity. Electric charges, positive or negative, exert electric forces on each other. Moving charges (electric currents) exert magnetic forces on each other. At first, this also looks like action at a distance. But the English physicist Michael

Faraday had a different idea; he said that there are electric and magnetic fields in space.

- o What is the electric field? At every point in space—even empty space—the electric field is present and has a certain strength and a certain direction. The field might vary from place to place, and it might change with time. Electric charges are the sources of the field. They determine the field values everywhere in the universe.

- o Given that the field exists, the electric force on any particular charge is completely determined by the local field. That is, an electrically charged particle gets its "marching orders" locally, from the field at the point where the charge itself is.

- o We can express this as an equation. If the local value of the electric field is E and we put an electric charge q in the field at that point, then the force, F, experienced by the charge is given by (qE).

- o This means that the same field can affect different particles in different ways. A positive charge is pushed in the same direction as E. A negative charge is pushed in the opposite direction. A neutral particle, such as one with $q = 0$, experiences no force at all.

- o We can see the effect of the field when we see charges pushed this way or that, but the field exists in space, even if no charges happen to be present.

- To sum up, Faraday's idea was that charges determine the arrangement of the field in space. The field determines how charges move through space. Magnetic fields and forces work in almost the same way.

The Gravitational Field

- It's natural to imagine that the same idea could be applied to gravity. Masses determine a gravitational field everywhere in space. That

gravitational field exists even in empty space. For example, the Earth sets up a gravitational field in this room. That field is nearly uniform, and it points downward. An object, such as an apple, experiences a force due to the local gravitational field.

- Much of gravitational physics sounds quite natural in this field language. For instance, the principle of equivalence in field language states: All objects have the same acceleration in a given gravitational field. A statement of the tidal effect in field language would be: The gravitational field can have different values in different places. A statement of hydrostatic equilibrium would be: At each point in a body at equilibrium, pressure forces balance the local gravitational field.

- We can also identify several new concepts in this thinking about fields:
  - The field exists even in empty space.

  - The field is not the same thing as the force it exerts. The same field may exert different forces on different objects in the field.

  - Objects get their marching orders locally in the field.

  - A field can vary from place to place and change over time, but it doesn't move. If mechanics is the science of force and motion and a field isn't a force and doesn't move, then a field is beyond mechanics.

Eliminating Action at a Distance

- According to Faraday, an object gets its marching orders locally. There's no instantaneous action at a distance. But doesn't the field itself get its marching orders at a distance?

- In the electric field, a body with electric charge is surrounded by a field. Faraday visualized this by drawing lines in space, with the direction of the lines indicating the direction of the field and the density of the lines indicating the strength of the field. The charge

is surrounded by radial field lines, meaning that the field is radial in direction at any point, and it's strongest closest to the charge when the field lines are bunched together.

- If we move the charge, the field does not respond instantaneously. The change in the field spreads outward at a finite speed.

- To describe this precisely, we need field equations; these are mathematical relations governing how the field varies in space and time. These field equations do not directly describe how forces affect motion; therefore, they are not really part of Newtonian mechanics.

- For electric and magnetic fields, the field equations are those put together by James Clerk Maxwell, a 19th-century Scottish physicist. Using his equations, Maxwell was able to show that changes in electric and magnetic fields move through empty space in the form of waves. The speed of the wave is determined by electric and magnetic physics.
    - The electric force between two charges is governed by an electric constant, something like Newton's constant for gravity (G). Another constant determines the magnetic force between two wires carrying electric current.

    - Maxwell showed that the speed of electromagnetic waves depends on these two constants. From the results of electric and magnetic experiments, he calculated that speed (c) as equal to 3×10^8 m/s, or about 300,000 km/s—the speed of light.

- According to Maxwell, light is an electromagnetic wave. It is a moving, oscillating disturbance of electric and magnetic fields in empty space.
    - Fields are not just complicated ways to talk about forces but are real in themselves. They have physical phenomena of their own without any particles of matter being present.

- Because light can carry energy and momentum from place to place, the fields can carry energy and momentum from place to place. But these fields are not part of mechanics. Mechanics is the science of force and motion—the science of how particles of matter move. Because fields are not made of matter, they lie outside Newtonian mechanics.

Einstein and Galileo

- The idea of thinking about electromagnetic fields in other than a mechanical way was difficult to accept in the 19th century. The universe was thought to be filled with a pervading substance called the aether. If that were true, then light waves would be mechanical disturbances, similar to sound waves; light would be vibrational waves in the aether. But Maxwell and others found it hard to make this idea work out in detail. Specifically, the fact that the speed of light seems to be the same in all directions was a puzzle.

- In the late 19th century, Ernst Mach, an Austrian physicist, wrote a critique of the basic concepts of Newton's mechanics. According to Mach, absolute space, absolute time, and intrinsic inertia were not justifiable as basic assumptions. He said that physics needed to be rebuilt on a more reliable foundation, setting the stage for the appearance of Albert Einstein.

- Recall that Galileo imagined doing various experiments in a smoothly moving ship and found that no experiment about force and motion could detect the movement of the ship. This is Galileo's principle of relativity; it explains why we cannot feel the motion of the Earth through space.
  - If the aether idea were correct, we could detect the ship's movement through the aether by doing an electromagnetic experiment, such as measuring the speed of electromagnetic waves—the speed of light—in different directions.

  - Einstein asked: What if the ship was moving at the speed of light? Then light could not move forward in the lab at all. That would be different from being at rest. The principle of relativity

would be wrong. To Einstein, that seemed too high a price to pay.

- Instead, he proposed a bold idea: Perhaps we cannot detect the motion of the ship even by speed-of-light measurements. We need to elevate the principle of relativity to a fundamental law governing all of physics, mechanics, electromagnetism, and beyond.

Ernst Mach made many discoveries in aerodynamics, but he also wrote a sharp critique of the basic concepts of Newton's mechanics: absolute space, absolute time, and intrinsic inertia.

- Einstein's first postulate is that every observer, whether at rest or in uniform motion, sees the same general laws of physics. The second postulate is that the speed of light, c, is the same, independent of the motion of the source. It follows that every observer, at rest or in uniform motion, must measure the same speed of light in every direction.

- The constant speed of light is impossible to reconcile with the aether theory, so we must eliminate that theory, but we retain Maxwell's equations about electric and magnetic fields. This idea is also a bit counterintuitive. Suppose I stand at rest and shoot a beam of light at speed c. You pass by me at speed $1/2c$ in the same direction. Common sense says that you observe the light beam traveling ahead of you at $1/2c$. But Einstein says that's not true; you would observe the same speed, c, for the light beam. How can this be true?

- Einstein's answer is profound: The world consists of space and time. Different observers measure space and time in different ways. These differences are exactly what make everything consistent.
  - Different observers actually agree about quite a lot: the general laws of physics, the speed of light in empty space, which events take place in the universe, and which objects are accelerating and which are not.

  - Different observers might not agree about which objects are moving and which are not, the time and place that various events occur, and the behavior of clocks.

- By holding onto Galileo's insight that we cannot define absolute motion, Einstein was led to a new framework for physics. As we'll see in the next lecture, in trying to understand gravity, Einstein reached back to another of Galileo's basic discoveries: that all objects in a gravitational field fall with the same acceleration. Einstein saw this familiar fact in a completely new way.

Suggested Reading

Einstein, "Ether and Relativity," in *The Tests of Time* (Dolling et al., eds.). Einstein ponders the physical meaning of relativity in this 1920 address at the University of Leyden.

Einstein and Infeld, *The Evolution of Physics*, chapters 2–3.

Faraday, "The Concept of the Electromagnetic Field," in *The Tests of Time* (Dolling et al., eds.). Michael Faraday, one of the originators of the field concept, explains its meaning and significance.

Michelson, "The Ether and Optical Experiments," in *The Tests of Time* (Dolling et al., eds.). A review by Michelson himself of the famous ether-drift experiments he did with Morley.

Questions to Consider

1. Is the field picture of electromagnetic forces (or gravitational forces) fundamentally simpler or more complicated? Explain your answer.

2. In 1905, Einstein laid out two postulates for his special theory of relativity. What would Galileo or Newton have thought of them? Imagine a dialogue among the three of them about the meaning of the principle of relativity.

The Falling Laboratory

Lecture 14

In the last lecture, we saw how Einstein elevated Galileo's relativity to a universal principle. The laws of physics are the same in every frame of reference. Whether an observer is at rest or in uniform motion, that observer sees the same physics, including the speed of light in a vacuum. This is called the special theory of relativity, "special" because it deals with observers in uniform motion. Unfortunately, the special theory of relativity excludes gravity. Electromagnetic and other forces fit beautifully into special relativity, but gravity is different. Einstein needed a new principle to guide him in thinking about gravity, and in 1907, he found it.

Frames of Reference

- Walking home from work one day, Einstein saw a building encased in scaffolding and wondered what a worker would experience if he were to fall off the scaffolding. We know that everything falls with the same acceleration (g), but for the worker—an observer in free fall—anything falling with him would not appear to accelerate at all; gravity would appear to be zero.

- Newton's principle of equivalence states that inertial mass is equivalent to gravitational mass. In free fall, an object, such as an apple, falling with the worker obeys the principle of inertia. If the falling worker pushes on the apple it accelerates, but without a push, it appears to move with constant velocity. A freely falling frame of reference, therefore, is an inertial frame of reference.

- According to Einstein, the frame of reference of the room you're in is not an inertial frame. If you let go of an apple in the room—exert no force on it—the apple accelerates downward. The force of gravity is acting on the apple. That force of gravity is just our way of saying that things in the room accelerate even if we don't push them. Einstein says that an inertial frame of reference is one

in which the law of inertia holds: If there is no force on an object, it does not accelerate.

- According to Einstein, gravity is just an effect of inertia. Gravity is really due to our choice of frame of reference. We can imagine two frames of reference: a freely falling frame and a frame tied to the Earth, such as your room. The inertial frame and the stationary frame are not the same thing, and the difference between them is the acceleration of gravity.

Implication of Einstein's Frame

- To sort out the implications of Einstein's idea, we'll consider a physics laboratory under four different conditions. In the first, the laboratory is at rest in deep space. Because it is millions of light-years from anything else, the gravitational field in the lab is zero. In an experiment in this lab, free bodies—bodies that are not being pushed—move in straight lines in the lab at a constant speed. The principle of inertia holds.

- In the second situation, the laboratory is at rest near the Earth. There is a gravitational field in the lab pointing toward the floor, so things in the lab accelerate downward at about 10 m/s^2. Not everything accelerates. You can stand still on the floor because the floor is exerting an upward force on you. In the absence of other forces, everything in the room accelerates the same. We say there is a force of gravity acting, and it's a universal force. It affects everything in the lab in exactly the same way.

- In the third situation, the laboratory is in free fall near the Earth. Because the lab is near the Earth, we would like to say that the gravitational field is not zero in the lab, but we would see no visible effects of gravity. Free bodies move in straight lines at a uniform speed. The principle of inertia holds.

- In the fourth situation, the laboratory is once again in deep space, millions of light-years from anything. We would like to say that the gravitational field is zero, but the lab is accelerating toward the

ceiling of the lab, pushed by a rocket motor below the floor, at 10 m/s^2. Because of inertia, objects not attached to the lab will tend to move at a uniform speed and direction in space. Those objects will appear to accelerate downward toward the floor at 10 m/s^2, and any object in the lab would appear to accelerate in exactly the same way.

- From outside this lab, we might say that isn't really acceleration but inertia. It's the laboratory that is actually accelerating, but in the frame of reference of the laboratory, everything looks exactly as if a uniform gravitational field exists inside the lab.

- If we were standing in the lab, we would have the sensation of weight. That sensation is just the force exerted by the floor on our feet by one part of our body on another part. In the accelerating laboratory, those forces are due to the rocket thrust that accelerates the lab, but the effect is indistinguishable from gravity.

- According to Einstein, the first and third situations are physically equivalent. No experiment in either lab can tell you whether you are in a region of space with no gravitational field or falling freely in a gravitational field. Both situations are true inertial frames of reference.

- Einstein also says that the second and third situations are physically equivalent. No experiment inside the lab can tell whether you are at rest in a gravitational field or accelerating through deep space. In one case, objects accelerate toward the floor because of Earth's gravity. In the other case, objects accelerate toward the floor because of inertia and because of the rocket motor attached to the lab. In either case, the observed acceleration of objects toward the floor is the same.

- Four hundred years ago, Galileo made two discoveries about motion: that the principle of inertia applies to all objects and that

all objects fall with the same acceleration due to gravity. Einstein essentially claims that those two facts are linked.

- Einstein is proposing a new principle of equivalence, not an equivalence between one kind of mass and another. For Einstein, there is only one kind of mass. This new equivalence is between apparently different physical situations. Free-fall and zero-gravity situations are physically equivalent. Gravitational and accelerated situations are physically equivalent. That is why gravitational and inertial mass are always exactly equal—because gravity and inertia are really the same thing.

Implications of Einstein's Equivalence

- Einstein's principle is a general principle of physics. Like the principle of relativity, all kinds of physical phenomena must be consistent with it. The equivalence we've drawn holds for any sort of experiment. For instance, it holds for an experiment about light.

- Imagine an accelerating lab in deep space. Acceleration defines up and down inside the lab; it has a floor and a ceiling.
    - We send a radio signal from a transmitter on the floor to a receiver on the ceiling. Of course, the signal takes some time to travel that distance.

    - During that time, the lab speeds up slightly as it accelerates. That means that by the time the signal reaches the receiver, the receiver is actually moving away relative to the transmitter when it sent the signal. As a result, there is a Doppler shift. The received radio frequency is slightly lower than the frequency that was transmitted.

    - The same thing should happen for any kind of electromagnetic wave, including light. The received light would be slightly redshifted.

- Now let's imagine a physically equivalent situation: a radio transmission in a stationary laboratory on Earth.
    - According to Einstein's principle of equivalence, we must observe the same phenomena we observed in the accelerated laboratory: The received signal at the ceiling must have a lower frequency.

    - In other words, gravity has an effect on waves, a gravitational Doppler effect. This must be true for light waves, as well. This is called gravitational redshift. Light that travels upward in a gravitational field is reduced in frequency. It's a little surprising to find that gravity affects light.

- There's a good way to think of this based on quantum physics. As Einstein and others realized, the energy of light comes in the form of discrete packets, "particles" of light called photons. The energy of a photon is related to the frequency of the light. The formula here is $E = hf$, where E is the photon energy, f is the light frequency, and h is a constant of nature called Planck's constant.
    - A photon has no mass. It always moves at the speed of light. We can think of the photon energy as a kind of KE. If a photon moves upward in a gravitational field, it makes sense that some of its KE is turned into gravitational PE. The photon does not slow down, but its energy decreases.

    - By Planck's equation, any decrease in photon energy means a decrease in light frequency.

- The phenomenon of gravitational redshift is very important. To see why, let's return to our radio transmitter and receiver.
    - A radio transmitter is an oscillating electric circuit, and that circuit is like a clock. One tick of the clock is one cycle of the wave. The receiver at the ceiling, therefore, sees the transmitter signal at a lower frequency. That means the receiver sees the transmitter clock at the floor running slower.

© iStockphoto/Thinkstock.

If GPS engineers ignored gravitational time dilation, then GPS data would be off by 10 km after just 24 hours; thus, engineers make software corrections to account for time dilation.

- In other words, gravitational redshift, the effect of gravity on light waves, implies an effect of gravity on time—gravitational time dilation. Clocks that are lower down in a gravitational field run slower. That relation between gravity and time is a major surprise.

- Gravitational redshift and time dilation have been observed in white dwarf stars and in the Pound-Rebka experiment on Earth. In this experiment, researchers determined extremely tiny shifts in gamma ray frequencies between a source of gamma rays and a detector about 22.5 m above the source. This frequency shift was about 2.5 parts per quadrillion. A clock in the same position as the gamma ray source would run that much slower than a clock in the position of the detector.

- Galileo observed that everything falls with the same acceleration. Newton told us that there are two quantities for an object, inertial mass and gravitational mass, and they always have the same value. According to Einstein, this means that gravity is not exactly a force at all. Friction, pressure, and electric and magnetic forces are all forces in the usual sense, but Einstein discovered that gravity is different. The frame of reference tied to the Earth, a stationary frame, is not the frame of reference in which the law of inertia holds, a free-fall frame.
    - What we call gravity is just that mismatch of the two, which means that gravity affects light and time.

    - These effects are tiny near the Earth, but the fact that gravitational redshift and time dilation occur is supremely important. It means that gravity distorts the very geometry of space and time.

    - To understand what that means, we need to think of space and time as one unified entity: spacetime, the four-dimensional world that is the realm of all physics.

Suggested Reading

Einstein and Infeld, *The Evolution of Physics*, chapter 3.

Wheeler, *A Journey into Gravity and Spacetime*, chapter 2.

Questions to Consider

1. You are a producer at a big Hollywood studio. Suppose a visionary director proposes the following: "Instead of merely simulating zero gravity in an airplane—as was done for the movie *Apollo 13*—we should launch a film crew into orbit and make a movie in *actual* zero gravity!" Respond to this (potentially *very* expensive) movie proposal.

2. Use the principle of equivalence to show that a horizontal light beam will be deflected downward in a gravitational field. (The bending of light by gravity is the subject of Lecture 18.)

Problem

Because the tidal effect is tiny over short distances, we may ignore it for our small laboratory. But suppose the lab was larger, and we had to take the tidal effect into account. Would the principle of equivalence still hold as we have stated it?

Spacetime in Zero Gravity

Lecture 15

In the 1600s, Galileo had discovered that everything falls with the same acceleration. In 1907, according to Einstein, that meant that free fall is equivalent to zero gravity and that gravity has laboratory effects exactly like an accelerating frame of reference. This is Einstein's new version of the principle of equivalence; it would lead to a new general theory of relativity. For this lecture, we will suppose that gravity is zero; we're working in a freely falling laboratory, whether that laboratory is near the Earth or out in deep space. Our frame of reference is inertial, and everything we say is covered by the special theory of relativity. In the next lecture, we'll see how to apply this lecture to gravity and Einstein's general theory.

Understanding Spacetime

- A mathematics professor named Hermann Minkowski came to an interesting conclusion about Einstein's work. He said that relativity is really a theory of four-dimensional geometry. To understand this idea, we need to make a detour through plane geometry.

- A plane is a two-dimensional surface. We often describe the plane using the coordinates x and y. The x axis is a horizontal line, and the y axis is a vertical line. Note that the two axes are perpendicular. Any point in the plane can be labeled by two numbers, x and y.
    - o The fact that it takes two numbers to locate a point on the plane is why we say the plane is two-dimensional. To locate a point in space, we would need three numbers, x, y, and z; space is three-dimensional.

    - o Describing points in the plane in this way goes back to the 17th-century philosopher and mathematician Rene Descartes. We can use such descriptions to relate geometrical ideas to numerical or algebraic relationships, as we see in an example of finding the distance between two points using the Pythagorean theorem.

- Minkowski asserted that the world we live in is really four-dimensional. It takes four numbers, three spatial coordinates and one time coordinate, to identify a point in space, for example, to arrange a rendezvous. In other words, it takes four coordinates to map our universe of space and time. Minkowski's term for this four-dimensional world was "spacetime."
    - Everything that happens in our universe lies somewhere in spacetime. The points in spacetime, things with both a where and a when, are called events.

    - The physical universe is the totality of all events. All of our statements about where and when, about length and duration, are statements about spacetime.

- A frame of reference is just a way of assigning space and time coordinates to events, something like a team of surveyors with measuring tapes and stopwatches finding three spatial coordinates and one time coordinate for each event. Different frames of reference—different teams of surveyors—might assign different space and time coordinates to events in spacetime, but the new coordinates describe the same spacetime, just as two teams of surveyors describe the same geography in mapping a county.

Thinking in Four Dimensions

- The first problem presented by Minkowski's idea is that it encompasses four dimensions, and most of us can't think in four dimensions. We simplify by imagining two or even one spatial dimension instead of three.
    - We can imagine one dimension of space, labeled by the x coordinate. We might imagine, for example, that everything we care about happens along a single line, like a long highway. With that dimension, we have one dimension of time, labeled by the t coordinate.

    - That lets us draw a picture of spacetime. The space coordinate, x, runs horizontally, and the time coordinate, t, runs vertically. The future is up and the past is down.

- Any event is just a point in this diagram. What about a physical object, such as an apple? At any moment in time, the apple has a location. If you put all these together, the apple is a long red line stretching from past to future. Minkowski calls this the “world line” of the apple.

- The shape of the world line in spacetime tells us how the apple is moving. An apple at rest will have a world line that is straight and vertical in our diagram. An apple moving right or left will have a world line that is tilted. A higher speed means more tilt. An apple that is accelerating will have a curved world line.

- Here’s the law of inertia in spacetime language: If there is no net external force acting on the apple, its world line is straight. That sounds simple, but it’s a fact worth remembering.

- Also notice that we can think of the time axis as a world line, too. It’s the world line for an object that’s at rest in the center of the spatial coordinates.

- Another problem with thinking in spacetime is that the spacetime axes have different units. Space is measured in m, and time is measured in s.
    - To address this problem, Minkowski multiplies time by *c*, the speed of light. In other words, the time coordinate is really *ct*; *c*, of course, is 300 million m/s.

    - Why use that speed? Because Einstein showed us that it’s a universal speed built into the laws of physics. Because the speed of light is so fast, even a short time is a long distance. One nanosecond (ns) is about 30 cm. One s is 300 million m.

Spacetime Interval

- Different frames of reference correspond to different coordinate axes, different ways of assigning four numbers, *ct*, *x*, *y*, and *z*, to events. Minkowski asked whether there is something that is the same in all frames of reference—something like distance—

even though the coordinates might be different. The answer is the spacetime interval, denoted by s. It's like a distance between spacetime points, and it's different in a surprising way.

- Let's suppose we're given the time and space coordinates for each of two events. The time interval between the events is Δt; we multiply that by c to get $c\Delta t$, the time difference measured in m. The spatial separation of the two events is Δx.

- The equation for spacetime interval is: $s^2 = c\Delta t^2 - \Delta x^2$. This looks similar to the Pythagorean distance formula, $d^2 = \Delta x^2 + \Delta y^2$, but in the spacetime interval equation, we see a minus sign rather than a plus sign.

- That minus sign is the difference between the familiar geometry of the plane and spacetime geometry.

• Here's Minkowski's essential idea: In different frames of references, different observers choose different space and time coordinates. They obtain different time and space intervals, $c\Delta t$ and Δx, between events. However, all observers will find the same spacetime interval between the same events as calculated by Minkowski's formula: $s^2 = c\Delta t^2 - \Delta x^2$.

• Let's look at an example: Event A happens at the origin, at coordinates 0, 0. Event B is 10 m of time later, or about 33 ns, and 8 m to the right; thus, $c\Delta t = 10$ m and $\Delta x = 8$ m. The spacetime interval between A and B is, therefore, given by $s^2 = 10^2 - 8^2$, which works out to 36; $s = 6$ m.

- That result looks a little strange on the spacetime diagram. Spacetime interval is like the length of the diagonal line from one event to the other. On our picture, it looks longer than either of the two sides, but it's actually shorter—only 6 m.

- How can that be right? The answer is simple: Our diagram is drawn in ordinary space, where the Pythagorean formula works. The events actually live in spacetime, where Minkowski's formula works.

Minkowski's Version of Einstein's Insight

- What if Δx is greater than $c\Delta t$? For instance, the two events may be simultaneous ($c\Delta t = 0$), but they happen at different points in space; they are separated only in space. Then, the formula tells us that s^2 is negative.

- That result is not as bad as it seems because spacetime interval is not exactly like distance. There is only one kind of distance between points in the plane, but there are three possible kinds of spacetime intervals between events: (1) timelike intervals, where time dominates and s^2 is positive; (2) null intervals, where $c\Delta t$ and Δx are the same size and s^2 is zero; and (3) spacelike intervals, where space dominates and s^2 is negative. All three of these have a slightly different meaning.

- Let's consider two events. The first is a flash of light (event F) from a camera flashbulb. The second is the flash of light reaching an observer standing 20 m away (event G). What is the spacetime interval between F and G?
    - Notice that on a spacetime diagram, the light flash traces out a diagonal line; F and G are separated in space by $\Delta x = 20$ m. The time separation is short because light travels very fast.

    - The universal speed c is $\Delta x/\Delta t$, and this means that the time separation measured in m is $c\Delta t$, which is Δx. In other times, the time separation between the two events is also 20 m. Because the time and space separations are the same, the spacetime interval, s^2, is zero. It's a null spacetime interval.

    - F and G, the origin and the destination of the light flash, are not the same event, but they are separated by a spacetime interval equal to zero. The reason for that is that light travels from one event to the other at the speed c.

- Consider the same story in a different frame of reference. We have different space and time coordinates, x' and ct'. The camera and the observer might be moving in this frame of reference. The distance

between them might be different. The values for $\Delta x'$ and $c\Delta t'$ may be different for the events in this frame; they no longer equal 20 m. But the spacetime interval is still null. Therefore, the new $\Delta x'$ and $c\Delta t'$ must still be the same size.

- o How fast does the light travel from F to G in this frame of reference? The speed in this frame of reference is: $v' = \Delta x'/\Delta t'$, which is equal to c.

- o The light flash travels at the same speed, c, in this frame of reference.

- Einstein phrases his insights as physics. The laws of physics, including the speed of light, are the same in different reference frames. Minkowski phrases his insights as geometry. Different coordinate systems in spacetime lead to the same spacetime interval between two events, just like the distance between two points in the plane. These two principles are really the same thing.

- At first, Einstein was a bit skeptical about Minkowski's idea, but he soon realized that the right way to think about physics is to think about spacetime. He realized that relativity is about spacetime relations between events. Minkowski's idea can also be used to verify time dilation, which Einstein had discovered in his original 1905 paper. Einstein would go on to use Minkowski's tools—spacetime, space and time coordinates, events, world lines, and so on—to build a new theory of gravity: the general theory of relativity.

Minkowski's ideas about the spacetime interval can be used to verify Einstein's discovery of time dilation—that a moving clock runs slower than a stationary one.

Suggested Reading

Minkowski, "Space and Time," in *The Principle of Relativity* (Lorentz et al.). The lecture in which Minkowski first introduced the idea of spacetime.

Taylor and Wheeler, *Spacetime Physics*, chapter 1.

Wald, *Space, Time and Gravity*, chapters 1–2.

Questions to Consider

1. Upon hearing the parable of the surveyors in our lecture, a friend says, "Why should we use coordinates at all? If the surveyors just measured and wrote down all the distances between points in the county, that would contain the same information—and no other surveyors would get different results." How would you answer your friend?

2. A nucleus of barium-137 is in an "excited" state (that is, a state with extra energy above the minimum). After a couple of minutes, the nucleus emits a gamma ray (a photon of light) and recoils a bit in the opposite direction. Draw a spacetime diagram for this process, showing the nucleus before and after and the photon. If you measure both time and space in m, at what angle should you draw the photon's world line?

Problem

A moving clock emits two flashes of light. When the clock is seen at rest, the two flashes are emitted just 12 m of time (40 ns) apart. Now we observe the same clock moving so that the two flashes occur 5 m apart in space. How much time (in m and ns) separates the two flashes in this frame of reference? By what factor is the moving clock running slowly?

Spacetime Tells Matter How to Move

Lecture 16

As we've seen, Galileo's observation that everything falls with the same acceleration gave Einstein some new insights: A frame of reference in free fall is equivalent to a frame with zero gravity; an accelerating frame of reference is equivalent to a gravitational field. Einstein used these insights to derive gravitational redshift and time dilation. Minkowski reformulated Einstein's relativity in terms of spacetime, the four-dimensional world that is the arena for all physics, and Einstein came to believe that spacetime is the best way to think about relativity. As we'll see in this lecture, Einstein then used Minkowski's spacetime tools to describe a strange new idea: that gravity is a kind of warp or curvature of spacetime.

Spherical Geometry

- Let's begin by returning to the geometry of the plane, specifically, straight lines. Given two points in the plane, 1 and 2, there is a line segment that is the shortest path between them.
    - Here's a simple way we might see this: We can choose x and y coordinates as we wish. Both points lie on the x axis.

    - The path from 1 to 2 must cover the whole distance along the x axis. A path that curves out of the x axis must cover a little more distance. The shortest path is one that lies entirely along the x axis—the straight path.

- We know what a point on a sphere is, but what is a straight line between points on a sphere? The simplest answer is that a straight line on a sphere is the shortest path between the points. This kind of path is called a great circle.
    - We can visualize a great circle as the intersection of the sphere with a plane that passes through the center of the sphere. On the Earth the equator is a great circle, but the other lines of latitude are not. The lines of longitude are great half-circles running from the North Pole to the South Pole and so on.

- The shortest path between two points on a sphere is a section of a great circle. If you stretch a piece of string tightly between two points on a globe, the string follows a great-circle path between those two points. It follows a straight path along the spherical surface.

- This path is also straight in another sense. Suppose you were to walk on the sphere, and so that you do not turn right or left, you make sure that the steps taken by your right foot equal the steps taken by your left foot. Your path will be a section of a great circle. If you want to walk in a straight line on a sphere, you don't have to take your marching orders from distant points at the beginning and end of your journey; you just follow your nose.

- We now know what a straight line would be on the spherical surface of the Earth, but these straight lines can look surprising on a map. As cartographers know, it's impossible to make a perfect flat map of the curved Earth without some distortion. For a small region on the Earth, that distortion is tiny, and we can ignore it. But if we try to map a larger region of the Earth or even the whole surface of the Earth, then map distortion is inescapable.
  - On a Mercator projection, an inch (representing 100 miles) at the equator represents only 50 miles at 60˚ north or south latitude. The farther you go from the equator, north or south, the worse that distortion becomes. At 80˚ latitude, near the pole, that inch represents only about 17 miles. The lesson is that east-west distances are shorter than they appear at points far from the equator.

  - Airline routes between cities are approximately great circles, but they appear curved on a map. Seattle and Paris are at about the same latitude, but flying directly east from Seattle is not the shortest path to reach Paris. The shortest path is a great circle that bends far to the north, flying over Greenland.

© Hemera/Thinkstock.

Airlines try to conserve fuel by flying the shortest paths between cities; thus, their routes are approximately great circles.

- Remember, east-west distances are much shorter than they appear on the map when we are close to the pole. The ideal path of the airplane makes use of this "arctic shortcut effect." It's worth going a little way north to take advantage of the shortcut.

- A German mathematician, Bernhard Riemann, made this spherical geometry far more general.
  - Riemann's geometry applies not only to a sphere but to a surface of any shape. It applies not only to a two-dimensional surface but to a space of any number of dimensions.

  - A straight line in Riemann's geometry is called a geodesic. A geodesic is the shortest path between two points in a curved space.

- Riemann's curved geometry is intrinsic; that is, we do not need to think about the space from the outside. The geometry of space is determined by distances between points within the space.

Minkowski's Spacetime and Riemann's Geometry

- For a general theory of relativity, Einstein found that he needed to combine Minkowski's spacetime with Riemann's curvature. Let's return to Minkowski's spacetime and draw a spacetime diagram. We have two events, labeled 1 and 2. We have a world line, a possible spacetime path, for an object passing between them.

- We can change the coordinates—shift our frame of reference so that events 1 and 2 occur at the same point in space; the two events differ only in time. Let's imagine various world lines joining 1 and 2.
  - A straight world line represents an object at rest. A curved world line represents an object that moves back and forth, starting at 1 and ending at 2.

  - We can think of each world line as the world line of a clock. Remember that a moving clock runs slower; therefore, the time measured by a clock moving from 1 to 2 is greatest if that clock stays at rest, if it follows the straight world line.

- In spacetime, a straight world line is a path of longest clock time between two events.
  - We chose coordinates so that events 1 and 2 occurred at the same spatial location, but that was just a choice of coordinates. The fact that the straight world line is a path of longest clock time remains true even if 1 and 2 were not at the same spot. A clock will register the longest elapsed time going from 1 to 2 if it follows that straight path. If it follows some other path, the clock will register a somewhat shorter time.

  - Straight world lines are particularly important for physics because of the law of inertia. In spacetime, if no net force acts on an object, then its world line is straight. That is, it has a

constant-velocity path. The law of inertia is where physics meets spacetime geometry.

- Remember Einstein's essential insight about gravity: A freely falling reference frame is an inertial frame. In free fall, an object that is not pushed will move in a straight-line path. A frame of reference attached to the Earth is not inertial. In that reference frame, objects appear to accelerate even if they're not pushed. That's what we've been calling gravity. The world line of an object, in other words, is what curves.

- If you drop an apple in the room you're in, the apple accelerates downward. That's what we call gravity. If we draw this motion in a spacetime diagram, its world line appears only slightly curved. But according to Einstein, the world line of the falling apple is not curved at all. It looks curved in our diagram, but that diagram is based on our frame of reference attached to the Earth. The curvature is a result of our map projection. The world line is a geodesic path in this curved piece of spacetime.

- What is a geodesic path in spacetime; what is a straight world line? As we've seen, a straight world line is the world line of longest clock time. Consider two events: The apple drops and the apple is caught down below. There are many possible world lines connecting the two events, each describing some possible motion of the apple.
    - Einstein claims that if we attach a clock to the apple, then that clock will register the longest elapsed time between start to finish if the path of the apple is the free fall path, that is, if the path is a gravitational trajectory.

    - That path is the straightest path, a geodesic in the curved spacetime near the Earth.

- Does this result make sense? If we want to maximize clock time, wouldn't it be best to move uniformly from the top to the bottom? Wouldn't the straight-line path in this sense look straight in our diagram? That's what we found in Minkowski's spacetime.

But gravity affects time. In a gravitational field, lower clocks run slower and higher clocks run faster.

- Consider two vertical segments in our diagram of the region of spacetime near the Earth. The two vertical segments each represent a certain amount of time, and they appear to be the same length; there's the same change in the *ct* coordinate. But the outer segment is really a longer physical time because a real clock runs faster farther from the Earth.
    - If our job is to find the best path for the apple from start to finish—the geodesic path, the path of longest clock time—then we should spend a little extra time farther out, where clocks naturally run faster. We'll cover most of the distance downward in the later part of the path. That's the path that's taking advantage of gravitational time dilation to maximize the elapsed time on the apple's clock.

    - Therefore, the geodesic path will appear curved. If it follows the geodesic path, the apple will appear to accelerate downward in the gravitational field. Why does that straight path appear curved? The answer is found in the map projection problem.

    - An airplane flying from Seattle to Paris seems to follow a curved path on the map, but the pilot does not turn the aircraft to follow it. The route looks curved because we aren't taking into account actual physical distance. East-west distances are shorter in the north, so the airplane can follow a shorter path by making some of the trip up north.

    - Similarly, the world line of the apple appears curved as it falls, but that's just our map projection for spacetime. The apple does not "feel" as if it's accelerating at all; it's just following an inertial path. The world line looks curved because we aren't taking into account actual physical time. Clocks run a bit faster farther from the Earth, so those times in the diagram are longer. The apple can follow a path of longer time by spending more of its trip farther out.

- What is special about the path the apple follows? In spacetime geometry, it's the straightest world line, and the same path would be the straightest path for any object. It's only the shape of the path that matters, and that's why all objects fall in the same way in a gravitational field.

- We've described different motion in gravitational fields—free fall, projectile motion, orbital motion—as paths in space shaped by gravitational forces. But for Einstein, all those kinds of paths are really geodesic paths, the straightest paths in curved spacetime.

The Effect of Gravity on Time

- From the principle of equivalence, we inferred that gravity must affect time. This gravitational time dilation is an effect that's very slight near the Earth, but it's a key factor in shaping the geodesic path of the apple in this lecture. Gravitational time dilation keeps the straight path, the inertial path, from being simple downward motion at constant velocity.

- We have said that gravity has a certain effect on time. In reality, what we call gravity is that effect on time. The effect of gravity on time is not a side issue or a curiosity but the heart of the matter. Gravity is curvature of both space and time, but the world lines we usually encounter are much longer in the time direction. It's the time effect that is important for the gravity that we know.

- We've explored Einstein's idea that gravity is not a force but a curvature of spacetime. To understand this, we combined the geometrical ideas of Minkowski and Riemann. The trajectories of free-fall bodies are spacetime geodesics, and that's true for falling apples and orbiting satellites. Those trajectories appear curved but only because of our map projection. The key ingredient in determining those trajectories is gravitational time dilation. The gravitational field around any star or planet is really a field of time distortion.

Suggested Reading

Taylor and Wheeler, *Spacetime Physics*, chapter 3.

Wheeler, *A Journey into Gravity and Spacetime*, chapter 10.

Questions to Consider

1. We defined a straight line on a spherical surface in two ways: as a path of minimum length and as the path followed by someone walking on the surface without turning right or left. Which of these concepts is more like the field idea, in which a moving body gets its marching orders locally? Try to describe great-circle paths on a Mercator map by means of a field that bends and stretches the path.

2. Gravitational time dilation, which we introduced in Lecture 14, is extremely small near the Earth's surface. Nevertheless, we now claim that this phenomenon is responsible for the action of gravity on a freely falling object, such as a projectile. How can such a tiny physical cause give rise to such an apparently large effect?

Matter Tells Spacetime How to Curve

Lecture 17

As we've seen, gravity is not a force but a warping of spacetime. A body in free fall steered only by gravity follows a geodesic. That path appears curved, but it is an inertial path, a path of the longest clock time between events. As John Wheeler, an American gravitational physicist, put it: "Spacetime tells matter how to move." This statement is true, but it's only half the story. By joining Riemann's curved geometry and Minkowski's spacetime, Einstein found a new way to think about gravity, but where does spacetime curvature come from? That's the second half of Wheeler's aphorism: "Matter tells spacetime how to curve." To make his theory of gravity complete, Einstein had to figure out how matter communicates that information.

Thinking about Curvature

- We can identify two kinds of curvature on two-dimensional surfaces: positive curvature, such as on the surface of a sphere, and negative curvature, such as on the surface of a saddle or the side of a vase.

- One way to see how geometry is different on a curved surface is to consider triangles. In Euclidean geometry, the angles of any triangle add up to exactly 180˚. In spherical geometry, in a surface of positive curvature, the angles of a triangle always add up to more than 180˚.
  - Consider a triangle on a globe consisting of one pole and two points on the equator. Each of the equatorial angles of the triangle is automatically 90˚; thus, it's easy to make a triangle with three right angles. That adds up to 270˚.

  - In saddle geometry, in a surface of negative curvature, the angles of a triangle always add up to less than 180˚.

- Another way of defining curvature is based on parallel lines. In Euclidean geometry, parallel lines do not get closer together or farther apart, but in spherical geometry, parallel lines tend to converge. In saddle geometry, parallel lines diverge.

- We use the deviation equation to define the amount of curvature as a measure of the tendency of parallel lines to converge or diverge. This equation is $a = -\mathcal{R}v^2d$, in which d is the distance between two particles, v is the speed at which they move along initially parallel lines, a is their relative acceleration (positive as the particles tend to diverge and negative if they converge), and $\mathcal{R}$ represents the curvature (positive if a is negative and negative if a is positive).

Translating Curvature to Spacetime

- Imagine that the parallel lines we're discussing are the world lines of particles. Straight lines are geodesics, the world lines of freely falling particles. How do we make two initially parallel world lines? We start with two particles close together in space; the particles are both initially at rest in the same frame of reference. Their world lines, then, will be parallel in a spacetime diagram.

- Consider this example: We take two apples and drop them at the same moment. Initially, both apples are at rest. Their world lines are parallel. Both apples follow geodesic paths in spacetime as long as they are in free fall.

- Is there any tendency for the apples to get closer together or farther apart? That would be a sign of curvature in spacetime. The naïve answer is no. The apples fall with the same acceleration; therefore, they stay the same distance apart, and there is no curvature. But as gravitational physicists, we know that there is a tendency for the apples to move together or apart; that tendency is a tiny tidal effect.
  - In Lecture 8, we learned that the force of gravity on the apple is not quite the same at every point. For one thing, the force of gravity always points toward the center of the Earth, so it lies in a different direction in different places. The force of gravity is weaker if the apple is farther away.

- This leads us to Einstein's realization: Gravitational acceleration by itself is no indicator of spacetime curvature. If you adopt a freely falling reference frame, then there is no acceleration. It's just an issue of map projection. The tidal effect is the sign of curvature. Here, we're taking a Newtonian concept and seeing it through Einsteinian eyes.

- Consider our two apples falling near the Earth, one above the other. The upper apple accelerates less. The lower one accelerates more. The apples tend to separate. There is a positive relative acceleration. These parallel world lines, therefore, tend to diverge from each other. That means a negative curvature in spacetime.

- Suppose the two apples are side by side, separated by the same distance as before. Each one accelerates toward the center of the Earth. That means they are also slightly accelerating toward each other. The relative acceleration in this case is just half as much as in the upper and lower example. These parallel world lines tend to approach each other. That's a sign of positive curvature in spacetime.

- Thus, spacetime is curved. The curvature is related to the familiar tidal effect. That is, it's related to differences in gravitational acceleration at different points. We can tell from our example that the curvature can be different in different directions.

Calculating Spacetime Curvature

- We can calculate spacetime curvature using the tidal effect equation from Lecture 8. If we have two apples separated by a vertical distance, d, and they're at a distance r from the center of the Earth, which has a mass M, the relative acceleration of the two falling bodies is given by $a = \dfrac{2GMd}{r^3}$. A positive value here means that the apples are drawing apart in space.

- Next, we apply the deviation equation: $a = -\mathcal{R}v^2d$. What do we mean by v here? The particles start out at rest, but they're still, in a sense, moving through time. Each s, they go 300 million m of time into the future. That's a spacetime speed of c, the speed of light. Putting the deviation equation together with the tidal equation, we get $-\mathcal{R}c^2d = \frac{2GMd}{r^3}$. The d's cancel, and we can solve for the curvature: $\mathcal{R} = -\frac{2GM}{c^2r^3}$. This is the spacetime curvature in the vertical direction.
- In the horizontal direction, the apples converge, and the relative acceleration is half as great. For spacetime curvature in the horizontal direction: $\mathcal{R} = +\frac{GM}{c^2r^3}$.
- Imagine six apples arranged around a central point: a top apple and a bottom apple, right and left, and front and back. All of them fall from rest, and each opposite pair of apples experiences a relative acceleration.
    - Each pair yields spacetime curvature in a different direction. The top-bottom pair yields negative spacetime curvature. The right-left and front-back pairs yield positive spacetime curvature exactly half as large. If we add up the curvature for all three perpendicular directions, the total is zero.
    - In an empty space, $\mathcal{R}_{\text{total}}$ is always zero. If the apples tend to approach in some directions, they tend to move apart in others. If spacetime curvature is positive in some directions, it is negative in others.
- What if there is some mass in the middle of the six apples? Then, the gravitational pull of that mass will pull the apples inward from all sides. Each pair of apples will approach each other. The total curvature, $\mathcal{R}_{\text{total}}$, will be positive; how positive depends on the mass between them.
    - To calculate the value for R_{total}, we calculate the rate of inward acceleration for pairs of opposite apples and find the

spacetime curvature. We then add up the curvatures of all three perpendicular directions.

- The answer is: $\mathcal{R}_{\text{total}} = \left(\dfrac{4\pi G}{c^2}\right)\rho$. The total spacetime curvature is equal to a constant multiplied by the density of matter (ρ). The density of matter at a point determines the total spacetime curvature there; matter tells spacetime how to curve.

- The constant in the equation involves G, Newton's gravitational constant; therefore, G determines how strong gravity is, that is, how much curvature a given density of matter produces. In ordinary units, G is tiny and c^2 in the denominator is huge. For ordinary densities of matter, spacetime curvature is very small.

Einstein's Equation

- What we have derived is a simplified form of Einstein's equation for the gravitational field. In Einstein's version, the right-hand side represents matter and energy at a point in spacetime.
  - One of the most basic laws of physics is that energy is conserved. It can move around and change form, but it is conserved.

  - We know, of course, that mass is equivalent to energy. That fact means that the divergence of the t quantity in Einstein's equation is zero. The total mass energy that goes into a spacetime point equals the total mass energy that comes out. No mass or energy is created or destroyed at that event. That's the conservation of mass energy.

- Amazingly, the left-hand side of the Einstein equation—the side involving curvature—must also have divergence equal to zero. This is a necessary truth of Riemannian geometry. We could not imagine a world in which that divergence is not zero. If the Einstein equation holds, both sides must have divergence equal to zero.

- At first, it might seem as if we have two independent laws of physics: the conservation of mass energy and Einstein's equation,

The statue of Einstein at the National Academy of Sciences shows him holding a notebook with three equations written in it: the equation for the photoelectric effect, $E = mc^2$, and the equation that relates mass and energy to the curvature of spacetime.

that is, the law of gravity. Logically, once we have the law of gravity, the conservation of energy follows automatically. It's not a separate principle. Put another way, Einstein's equation tells us what it is about the universe that ensures the conservation of energy. Gravity—spacetime curvature—enforces the conservation law.

- Einstein's quest began with a simple thought experiment: What would gravity be like for someone in free fall? The answer is: like no gravity at all. From that, we've been led with Einstein to a new kind of theory, a theory unlike anything that had ever been proposed before. Gravity is spacetime curvature. Spacetime tells matter how to move. Matter tells spacetime how to curve.

Suggested Reading

Wald, *Space, Time and Gravity*, chapter 3.

Wheeler, *A Journey into Gravity and Spacetime*, chapter 5. Wheeler explains the connection between tides and curvature in his own unique style.

Questions to Consider

1. We defined a surface of positive curvature as one in which the angles of a triangle add up to more than 180°. We also said that in such a surface, lines that are initially parallel eventually converge. Draw a diagram that shows how these two geometrical facts are connected. (Hint: The parallel lines lie along two sides of a triangle.)

2. In the classic science fiction novel *First Men in the Moon*, written in 1901 by H. G. Wells, the scientist Cavor invents a substance called "cavorite" that is "opaque to gravity." Anything placed within a shell of cavorite is no longer affected by the Earth's gravitational pull and, thus, flies in a straight line into space. (This is how the protagonists travel to the Moon.) Comment on cavorite in light of Einstein's general theory of relativity.

Problem

In the lecture, we said that we could derive a version of Einstein's equation from what we already know. Do this by imagining six apples on the top, bottom, forward, back, left, and right sides of a sphere of radius r that contains mass of density ρ . Here are the equations you need:

- Acceleration of each apple toward center of sphere: $a = \dfrac{GM}{r^2}$
- Mass of sphere: $M = V\rho$
- Volume of sphere: $V = \dfrac{4\pi}{3} r^3$
- Relation of curvature to relative acceleration: $a = \mathcal{R}c^2 d$

Light in Curved Spacetime

Lecture 18

We have been following Einstein on a journey to an amazing destination: general relativity, the theory that gravity is actually the curvature of spacetime. Spacetime is the arena of all of physics, which means that spacetime curvature affects all of physics, including time, distance, motion, and light. To illustrate how gravity and spacetime curvature affect absolutely everything, in this lecture, we'll see how gravity can change our thinking about something ordinary: temperature. How hot or cold an object is would seem to have nothing to do with gravity, but the temperature of an object does have something to do with light.

Thermodynamic Changes

- Imagine two walls facing each other with nothing in between; because both walls have a temperature, they emit photons of thermal radiation, which are absorbed by the opposite wall. In an equilibrium, each wall emits the same amount of energy it absorbs per second; thus, it is neither cooling down nor heating up. There will be equilibrium when the two walls are at the same temperature. This is a basic rule of thermodynamics: If two bodies can exchange energy and are in equilibrium with each other, then those two bodies are at the same temperature.

- Let's add gravity to this picture. Instead of two walls side by side, imagine one wall above the other. In other words, we have a floor and a ceiling in a gravitational field. The floor is at some temperature, emitting thermal radiation. The photons from the floor fly upward and lose energy as they go.
    - The floor emits some huge number of photons per second, but an observer on the ceiling sees everything on the floor happening slowly. That's gravitational time dilation. The radiation that reaches the ceiling has fewer photons per second, and those photons have less energy on average.

    - The reverse happens for photons going downward. They increase in energy as they fall. An observer on the floor sees everything on the ceiling happening quicker. That's the other side of gravitational time dilation. The photons arrive at the floor at a higher rate and with a higher energy on average.

    - Therefore, if the floor and the ceiling are to be in equilibrium—each one emitting the same amount of energy that it is absorbing per second—then the floor and the ceiling cannot be at the same temperature. The floor must be warmer and the ceiling cooler to compensate for the effect of gravity.

- The rule of thermodynamics that states that bodies in equilibrium must have the same temperature must be changed in the presence of gravity. If two bodies in a gravitational field are in equilibrium, then the lower one is at a higher temperature and the higher one is at a lower temperature.

Confirmation of Einstein's Theory

- In 1911, still several years before he completed his theory, Einstein predicted that light rays should be bent by the gravity of the Sun. He based this prediction on Newton's gravity plus the equivalence principle.

- Suppose we could observe a star right next to the visible disk of the Sun. The ray of light from the star is bent slightly toward the Sun as it passes; thus, we would see light coming from an apparent direction that is shifted outward from the Sun. Einstein calculated that the apparent position of the star should be shifted outward from the Sun by 0.9 arcsec.

- In 1915, with the full theory available, Einstein revisited this problem and realized that his 1911 calculation was not quite right because it didn't take into account the full curvature of spacetime. The actual deflection of light should be about 1.75 arcsec.

- Arthur Eddington, a British astrophysicist, realized that he could test Einstein's new theory by making a careful telescopic observation during an eclipse to measure the apparent positions of stars near the edge of the Sun's disk. After making such measurements in 1919, Eddington announced that Einstein was correct. The stars shift by 1.75 arcsec as the Sun passes.

- Eddington's observation was a powerful confirmation of Einstein's theory. Light is deflected by gravity. Further, the amount of the deflection agrees with Einstein's spacetime curvature theory, not simply Newtonian gravity.

Gravitational Lensing

- The bending of light by gravity is more than a curiosity. Recall Fritz Zwicky, the astrophysicist who realized that galaxy clusters contain more matter than we can see. In 1937, he also suggested that it might be possible to use the gravity of a galaxy or a galaxy cluster as a kind of cosmic lens.
  - Think of a convex lens made of glass. Light refraction and the shape of the lens cause light rays to be deflected inward toward the axis.

  - If we look through the lens, this effect magnifies objects behind it. If the lens has an irregular shape, it might form multiple images that might be stretched or distorted in some way.

- Consider a nearby galaxy with more distant galaxies behind it. Because of the gravity of the nearby galaxy, light from the distant galaxies is deflected inward, just as in a convex lens. Things behind the gravity appear slightly magnified.
  - The nearby galaxy might block a direct view of what's behind it, but to each side, the distant galaxies will appear shifted outward and slightly distorted. A round galaxy would appear as an oval or even a thin arc.

    - Because the mass of the nearby galaxy might also be distributed in an irregular way, we might even see more than one image of a single background galaxy.

- Like so many of Zwicky's ideas, this one took decades to bear fruit, but we now know hundreds of examples of the gravitational lens effect. An image from the Hubble telescope, for example, shows a distant cluster of galaxies about 2 billion light-years away, but the gravity of that cluster has warped and distorted the light of even more distant galaxies behind it. We see those galaxies as streaks and arcs around the periphery of the nearer galaxy cluster. Those distorted images are produced by gravitational bending of the light in the warped spacetime around the nearer cluster.

- We can use gravitational lensing to map out the mass that produces this effect, including the mass of both visible matter and dark matter.

The Speed Limit of the Universe

- Let's recall Minkowski's picture of spacetime from Lecture 15. We have a vertical axis, the *ct* coordinate, time measured in meters. We also have two horizontal axes, *x* and *y*. Points in spacetime are events. Physical objects are represented by world lines stretching from the past to the future.

- Suppose we send a light pulse from one end of a lab to the other. In spacetime, the world line of a photon is drawn at 45° because the light moves 1 m of space per m of time. Imagine that a light flash occurs at some event, and the light from the flash spreads out in all directions from the event. In spacetime, that light looks like a stack of circles of increasing radius—a light cone made up of 45˚ lines.

- Between the emission and absorption of a pulse of light, the spacetime interval is zero; it's a null interval. That's because the squared interval, s^2, is $c\Delta t^2 - \Delta x^2$, and that expression is zero for the emission and absorption events. We say that the world line of a photon is a null line, and the light cone of an event is made of null lines.

- An object at rest is a vertical line in our spacetime diagram, and two events on that world line—two ticks of the clock of an object—have a positive s^2 between them. We call that a timelike interval. If the interval is timelike in one frame of reference, then it must be timelike in any frame. That is, even if the object is moving, the spacetime interval between two points on its world line is always timelike. It can never be null; therefore, the object can never be moving at the speed of light.

- This is a famous fact of special relativity. No object that can be at rest can move at the speed of light; the speed c, 300 million m/s, is an absolute upper bound to the speed of ordinary objects. A photon is a slight exception. It always moves at speed c in a vacuum, never slower and never faster. There is a speed limit to the universe, and that speed limit is c.

Spacetime and Causal Structure

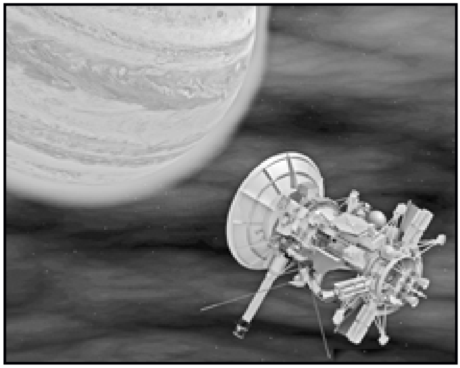
- Let's pick an event on the world line of an object in spacetime and call it here-and-now. The future world line of the object must lie inside the light cone of here-and-now. The light cone encloses every part of spacetime that the object could possibly reach. This is true for objects, all kinds of energy, and any sort of signal or information. The only places and times that can be reached from the here-and-now are those parts of spacetime that lie inside the light cone formed by a flash of light here-and-now.

- Physicists sometimes say that light cones describe the causal structure of spacetime. In other words, if you do something here-and-now, then the only things that you can affect by your action lie inside the light cone of here-and-now.
    - Alpha Centauri is a star 4.3 light-years away; that is, it would take 4.3 years to reach Alpha Centauri even moving at the speed of light.

    - Nothing that you do here-and-now can have any effect there 1, 2, or 3 years from now. Those events are as unreachable as

the past. They are, in fact, unreachable in almost the same way because they are beyond your ability to affect in any way.

- Because gravity affects light, it must also affect light cones. That is, spacetime curvature also warps the structure of possible cause and effect.
    - To understand this idea, let's use Eddington's coordinates—his map projection—for curved spacetime. In this diagram, roughly speaking, time is vertical and space is horizontal. A massive object, such as a star, appears as a thick world line running up the middle.



© iStockphoto/Thinkstock.

A radio time delay with the Cassini space probe provides an excellent and precise test of Einstein's general relativity within our own solar system.

    - Far away from the star, gravity is very weak. Light cones are symmetric, with 45˚ sides. Nearer to the star, in the region where the gravity is strong, the inside edge of the light cone, the side nearer the star, is still at 45˚, but the outside edge is slightly more vertical, as if outward-moving light moves slower.

    - Light isn't really moving slower; that's an effect of the map projection. But in the Eddington projection, light cones near a massive object are a little narrower and slightly tilted inward. This qualitatively shows a real effect of spacetime curvature.

    - Suppose we try to bounce a radar signal off the surface of a star. The ingoing signal travels along a 45˚ line, but the signal

does not return on a 45° line. The light cones are distorted in curved spacetime, and the light comes back slower.

- In Eddington coordinates, the spacetime curvature near a massive body produces a distortion of the light cones. As we will see in Lecture 20, when gravitational curvature of spacetime becomes very strong in the collapsed stars known as black holes, the distortion of the light cones becomes extremely important.

Suggested Reading

Eddington, "The Bending of Light Rays," in *The Tests of Time* (Dolling et al., eds.). Eddington gives his own account of the eclipse observations and their significance.

Gates, *Einstein's Telescope*, chapters 4–6.

Moffatt, *Einstein and Eddington*. This slightly fictionalized movie deals with Einstein's long labor to understand gravity, and how Eddington dramatically confirmed his work, bringing it to the attention of the world. The social and political difficulties faced by the two men are also vividly portrayed.

Questions to Consider

1. It is a remarkable coincidence that our Moon is just the right size to cover the visible disk of the Sun during an eclipse. Could Eddington have made his observations if (a) the Moon were much smaller, or (b) the Moon were much larger?

2. Every point in spacetime has two light cones associated with it: one that expands outward toward the future and a reversed one that expands outward toward the past. Make a spacetime diagram of this. Events in the forward cone are effectively "in the future," and events in the backward cone are "in the past." Explain this idea. What about events that are in neither light cone?

Gravitomagnetism and Gravitational Waves

Lecture 19

In Newtonian physics, the gravity of a planet does not depend on whether or not the planet is rotating, and the only physical effects of gravity are the forces exerted on masses. In Einsteinian gravity, however, things are different. In this lecture, we will turn to electromagnetism for guidance in looking at new gravitational phenomena. Electromagnetic waves include visible light, radio waves, and X-rays, and Maxwell's theory of electromagnetic fields incorporates all these phenomena. Einstein's theory of curved spacetime is, in many ways, the gravitational equivalent of Maxwell's field theory. That motivates us to ask: Is there a gravitational version of magnetic forces? Can there be gravitational waves in empty space? As we'll see, the answer to both questions is yes.

Determining Rotation

- As we've said, gravity is really about inertia. In a freely falling frame of reference, the law of inertia holds. A free body moves in a straight line at a uniform speed; its world line in spacetime is a straight path. A frame of reference attached to the Earth's surface, on the other hand, is not inertial. A free body accelerates downward; its world line is a geodesic—a straight line in curved spacetime.

- Let's consider a rotating lab, which is not an inertial frame. In a rotating lab, we feel many forces due to inertia, including the centrifugal force and the Coriolis force, a sideways deflection of moving objects. Anything moving toward the rotation axis is deflected in the direction of the spin, and anything moving away from the rotation axis is deflected in the opposite way. These apparent forces are fairly familiar.

- You can determine whether you are in a rotating reference frame by checking to see whether the law of inertia holds. Is there a centrifugal force in your reference frame? Do your arms tend to go outward even if nothing is pulling them? If so, then you are

rotating. If you want to stop rotating, then you just adjust your motion to eliminate these rotational forces. The best tool for finding a nonrotating frame in a lab is a gyroscope.

- o A gyroscope is a heavy object that can spin about an axis. If we support it at its center of mass, then it will not precess around when we set it spinning. Its angular momentum will be conserved.

- o Since the angular momentum points along the rotation axis, then that axis will point in a fixed direction in space—a fixed direction in an inertial, nonrotating frame. We can therefore use a gyroscope to detect whether or not our lab is rotating.

- Another method for detecting rotation is to use a telescope. You point a telescope at a distant star. If the star stays put in your field of view, then you're not rotating, but if the star drifts to one side, you are.

- Both of these methods—use of a gyroscope or a telescope—give the same results. Ernst Mach, the Austrian physicist we met earlier, believed that there was a physical reason for the agreement of gyroscopes and telescopes in determining rotation. Mach thought that distant masses in the universe might determine inertia in the lab. Thus, the inertial nonrotating frame for the lab was somehow fixed by the distant stars. Mach guessed that this might be some kind of gravitational effect; his suggestion is often called Mach's principle.

- Do the gyroscope and telescope methods always agree about rotation when we're talking about general relativity? Somewhat surprisingly, the answer is no. This application of Einstein's theory was worked out by the Austrian physicists Josef Lense and Hans Thirring.

The Lense-Thirring Effect

- Let's imagine a gyroscope that sits inside a massive shell of matter. The shell is not rotating, but it's transparent. We also have a

telescope that points along the axis of the gyroscope. When nothing is rotating, the telescope always points through the shell at the same star. The telescope and the gyroscope tests agree.

- We now start the shell of matter rotating rapidly.
    - The gyroscope axis is still determined by inertia. That is, if you are inside the shell and you orient yourself according to the gyroscope, you will feel no centrifugal force. Your reference frame is inertial.

    - When you then look through the telescope that's attached to the gyroscope, you see distant stars in space moving across the field of view. According to the telescope method, you are rotating.

    - The gyroscope and telescope no longer agree. Inside the shell, if you stop your rotation according to the gyroscope, then the stars appear to go around. If you stop your rotation according to the telescope, then you feel your arms pulled outward by centrifugal force. There's something about the rotation of the massive shell that has made the difference.

- What we're seeing here is a gravitomagnetic effect, produced by the moving mass in the shell. The rotation of the shell drags the inertial frame inside by a very slight amount. That's why the Lense-Thirring effect is sometimes called frame dragging.

- Does this effect contradict Mach's principle? We can think of the effect as Mach's principle in action. According to Mach, the whole universe determines the inertia in the lab by perhaps something like a gravitational effect, but the shell is massive and close. The rotating shell has an exceptionally large vote on what the inertial frame inside will be. The shell's rotation will pull the inertial frame along with it just a bit.

- Something similar would happen if we put a gyroscope beside a rotating planet. The rotation of the planet would drag nearby inertial

frames around due to the Lense-Thirring gravitomagnetic effect. The gyroscope axis would precess slightly relative to the stars in the direction of the rotation of the planet. The Lense-Thirring effect was recently confirmed by a NASA satellite experiment known as Gravity Probe B.

- Some places in the universe, such as neutron stars, experience a much greater frame-dragging effect than we do on Earth. A neutron star might be twice the Sun's mass but only 10 km in radius. It rotates up to 1000 revolutions per s. The axis of a nearby gyroscope next to a spinning neutron star would precess relative to the stars many times around per s. Near a rotating black hole, the frame-dragging effect can be even greater.

Gravitational Waves

- Recall Wheeler's motto: Matter tells spacetime how to curve. Matter in motion, a rotating planet or star, curves spacetime in a different way. The rotation of a massive body imposes a kind of twist to spacetime geometry, dragging inertial frames along with the rotation.
    - Other kinds of motion have other effects on spacetime, including extraordinarily rapid and violent motions, such as stellar explosions or collisions.

    - As the matter in these events redistributes itself, the gravitational field changes; spacetime changes around the masses. Those changes in spacetime spread outward into space as gravitational waves.

- A gravitational wave is a traveling disturbance in the curvature of spacetime. The wave moves through space, but it is also a property of the space through which it moves. Within the wave, space is empty; it contains no matter at all. A gravitational wave is simply a moving change in the geometry of space.

- Gravitational waves move at *c*. One way we can detect gravitational waves is to use a pair of inertial masses. This is the idea behind the

Laser Interferometer Gravitational Wave Observatory (LIGO), the most ambitious effort yet to detect gravitational waves from space.

- o LIGO consists of two facilities, one in Louisiana and one in Washington State. In LIGO, massive mirrors hang from wires 4 km apart. The distance between them is measured by interferometry.

- o The apparatus is fantastically accurate. It can detect a change in that 4-km distance of less than the diameter of an atomic nucleus. But even with such accuracy, LIGO could only detect the most powerful sources of gravitational waves because gravity is a weak force, and gravitational waves themselves are expected to be incredibly weak.

- o LIGO has been running for several years now and has not yet detected gravitational waves, although physicists are still fairly certain that they exist.

Gravitational Waves and Energy

- Given that gravitational waves are made out of nothing but empty space, it doesn't seem that they could carry energy. But suppose the two hanging masses in our detector are connected by an elastic thread. As the distance between the masses changes, the thread is stretched and relaxed. Over time, friction will heat up the thread; that's energy that has been added to the system. Gravitational waves must, in some manner, carry energy.

- About 20,000 light-years away is a remarkable star system called PSR 1913+16, also called the binary pulsar. In this binary star system, both stars are neutron stars, and one is a pulsar. The pulsar rotates rapidly and emits a periodic radio pulse toward the Earth; that radio pulse comes about 17 times per s. This precisely timed radio signal allows the motion of the stars to be tracked with extreme precision.

- The orbit of these two stars is a very close one. At the closest, the stars are less than 1 million km apart. The orbital period of the two

stars is only 7.75 hours. Thus, we have two very massive objects orbiting each other very quickly.

- o We've said that mere rotation could not generate gravitational waves because the distribution of mass within the rotating object does not change with time. It's in a steady state.

- o But here, we have two discrete masses moving around each other through space. The configuration of those masses does change rapidly with time. It's a perfect system for generating gravitational waves.

- There should be ripples in spacetime geometry spreading out from the binary pulsar. At a distance of 20,000 light-years, these waves are too weak even for LIGO to detect. But gravitational waves carry energy, which should cause the binary pulsar system to lose energy over time, and we can calculate how much energy that is. In fact, the gravitational wave energy output from the binary pulsar system is about 2% of the total light energy output of our own Sun.

- As the system loses energy, the two neutron stars move into a closer orbit. They move faster, and their orbital period becomes shorter. It's getting shorter by 76 microseconds (μs) each year. That doesn't sound like much, but it's precisely measurable thanks to the pulsar. The stars are spiraling toward each other and will eventually collide.

- Here's the important point: The rate of energy loss from this system exactly agrees with the prediction of Einstein's general relativity. That's why we can be so confident that gravitational waves exist. We cannot detect them directly yet, but we can measure the energy loss in the binary pulsar. That loss can only be reasonably explained by the ripples in spacetime spreading out from the rapid orbital motion of the two neutron stars.

Suggested Reading

Collins, *Gravity's Shadow*. Collins's rather lengthy book is a remarkable look at the long experimental search for gravitational waves. Though the focus is on scientists and how scientific ideas develop, the book describes a great deal of gravitational wave physics along the way.

Thorne, *Black Holes and Time Warps*, chapter 10.

Wheeler, *A Journey into Gravity and Spacetime*, chapter 11.

Questions to Consider

1. A classic museum demonstration of the Earth's rotation is the Foucault pendulum, first exhibited by Leon Foucault in 1851. A large pendulum swings back and forth, but because the Earth is rotating, the plane of the swing slowly precesses around relative to the building. Would you classify this as a gyroscope method or a telescope method?

2. Suppose a very strong gravitational wave passes through the room in which you are sitting. What sort of damage would you expect?

Gravity's Horizon—Anatomy of a Black Hole

Lecture 20

In this lecture, we resume a story we began in Lecture 12. There, we talked about the internal structure and life history of a star. At each stage of the star's lifetime, it was in hydrostatic equilibrium—a stalemate between the inward pull of gravity and the outward push of pressure forces. For an ordinary star, such as our Sun, that internal pressure is due to extremely high internal temperatures maintained by nuclear reactions in the core, but the nuclear fuel eventually runs out. Such stars as our Sun may then become white dwarfs; more massive stars may become neutron stars. For still more massive stars, no final hydrostatic equilibrium is possible. Once their nuclear fuel runs out, gravitational collapse is complete, and they become black holes.

An Early Picture of Black Holes

- The idea of black holes goes back to the late 1700s and the writings of two men, John Michell in England and Pierre-Simon Laplace in France.
  - Both Mitchell and Laplace knew that light travels at a finite speed, c, and they both guessed that gravity affects light.
  - In Newtonian gravity, every star or planet has an escape speed, v_{escape}. But if a body is massive and compact enough, then its escape speed will be faster than c. In that case, light would not be able to escape from the surface of that body into deep space.
- From Lecture 6, we know the formula for the escape speed of a star of mass M and radius r: $\sqrt{\frac{2GM}{r}}$. For a star of a given mass, how small would the radius have to be for the escape speed to equal c? Setting $v_{\text{escape}} = c$ in our equation, the solution for r is: $\frac{2GM}{c^2}$.

If the star is smaller than this radius, light will not be able to escape from its surface.

- If we put the mass of the Sun into the equation, 2×10^{30} kg, we find that r is about 3 kg, which is a bit smaller than a neutron star. This is a simple calculation based on nothing but Newtonian gravity; nevertheless, it yields a correct answer. This is both convenient and somewhat dangerous. We need to be careful how we think about this because the whole conceptual framework of Newtonian gravity is wrong.

- As we said, escape speed is how fast you have to go to escape a star or planet without any additional push. The escape speed for the Earth's surface is about 11 km/s, but you could actually escape the Earth's surface without going that fast by applying a continual force.
    - Imagine that we have a ladder 1 million km high. At that distance, the Earth's escape speed is less than 1 km/s. You could escape at that speed and never have to travel close to 11 km/s.

    - This kind of picture of escaping at slower than the escape speed does not work in Einstein's gravity. Black holes are something very different from the dark stars that Michell and Laplace imagined.

The Schwarzschild Radius

- A truer picture of black holes began to emerge in 1916. In that year, a German astrophysicist named Karl Schwarzschild discovered a mathematical solution to the Einstein equations that describes the gravitational field of a pointlike mass.

- Schwarzschild noticed that something odd happens at the radius $\frac{2GM}{c^2}$, the radius we calculated by setting the escape speed equal to the speed of light. In this context, it's the radius at which the gravitational redshift becomes infinite. This radius is now called the Schwarzschild radius.

- In Eddington's map of Schwarzschild's spacetime—the gravitational field of a point mass—time is vertical and space is horizontal. The point mass is a vertical line in the center, and light cones show how a flash of light would spread out from various points.

- Far from the mass, the light cones are at 45˚. Closer to the mass, the inward side of a light cone is still at 45˚, but the outward side looks more nearly vertical. At the Schwarzschild radius, the outward side is completely vertical. Inside the Schwarzschild radius, both the inward and outward sides tilt inward.

- A flash of light from an event inside the Schwarzschild radius cannot go outward at all. All directions are effectively inward, toward the point mass. Events inside the Schwarzschild radius can never be seen by an outside observer. This is a black hole.

- The Schwarzschild radius of a black hole is called its event horizon. Events inside are beyond the horizon and cannot be seen by an outside observer. Note that nothing is really present at the event horizon. It's a mathematical surface formed by the outer edges of light cones that are just barely trapped by gravity.

- Remember, the world line of an object must always stay within its light cones. That's the spacetime way of saying that *c*, the speed of light, is the speed limit for everything. Therefore, any object that crosses the event horizon can never come out. All such objects must move inward.

- No mere force can prevent that inward movement. A force acts within spacetime. It curves the world line of an object, but the inward movement of an object within the event horizon is built into the very structure of the spacetime geometry within which it moves.

Anatomy of a Black Hole

- All the mass of a black hole is in a central point—the singularity—where the density and curvature become infinite. In a sense, the singularity is a kind of edge to the spacetime we know. Surrounding the singularity is the event horizon. It keeps the singularity hidden from view because the horizon is a one-way surface for energy and information.

- Outside the horizon, the spacetime curvature—the gravity of the black hole—influences nearby objects, but the attraction of the black hole is not inexorable any more than the Earth's gravity is inexorable.

- What is a black hole made of? In a sense, it isn't made of anything. All the ingredients that made it have disappeared and been squeezed into the singularity. The final state of the black hole is extremely simple. In the words of John Wheeler: "A black hole has no hair." Its properties, such as chemical composition, simply no longer exist in the universe.

- One property that remains is the angular momentum of the black hole, and that's because of the Lense-Thirring effect. Rotation affects the spacetime geometry around an object. We can still see the rotation in the gravitational field around the black hole.

Observing Black Holes

- We can observe black holes by their interaction with other matter near them. The matter falling into a black hole generally forms something called an accretion disk, which is very dense and rapidly moving. The matter in an accretion disk is extremely hot and laced with strong magnetic fields.

- The accretion disk is a bright but compact source of X-rays. Because of rapid, turbulent motions in the small accretion disk, the X-rays fluctuate rapidly in intensity. Those rapid changes are key.

- A very large object cannot change its apparent brightness all at once because light from different parts of the object travels different distances to reach us.

- Even if the object did change its brightness in an instant, we would see a more gradual change as the light reached us at different times. A rapidly changing source of light or of X-rays must be relatively small.

- In 1970, NASA launched the Uhuru satellite, an orbiting X-ray observatory. One of the astronomical objects it discovered is Cygnus X-1, which is a highly variable X-ray source in the constellation Cygnus. The X-ray intensity from Cygnus X-1 varies in less than 10 ms. The source must be smaller in size than the Earth.
  - Ordinary telescopes also found a blue star called HDE 226868 at the same point in space, but the blue star is not the source of the X-rays.

  - Using the Doppler effect, we can detect changes in the velocity of HDE 226868. The star must, therefore, be orbiting with an unseen companion body. From the velocity changes, the mass of the companion is found to be at least 6 solar masses. That's too large to be a white dwarf or a neutron star. The only possibility is that it's a black hole.

  - What we have here is an ordinary star and a black hole in a fairly close orbit, so close that the black hole draws mass from the visible star and forms an accretion disk. The hot, dense inner part of the accretion disk, only a few thousand km across, produces rapidly fluctuating X-rays.

- Very large black holes are found in the centers of galaxies. The Milky Way contains a few 100 billion stars plus gas, dust, dark matter, and so on. It is disk-shaped, a spiral galaxy, and we are about 30,000 light-years from the center. The center of our galaxy

is actually in the direction of the constellation Sagittarius; we can see it with infrared and radio telescopes.

- The galactic center is a complicated and violent place. Infrared and radio telescopes show that stars near the center are moving rapidly, indicating the gravity of a massive unseen object. This massive object is several million times the mass of our Sun—a supermassive black hole.

- Black holes are fairly common in the universe. One place to find a black hole is in what astronomers call an active galactic nucleus (AGN). Galaxies with AGNs have extremely bright central regions. There are radio emissions, X-rays, and so on. An AGN can be observed billions of light-years away. There are several types of AGNs, including radio galaxies, quasars, blazars, and so on.

- Smaller black holes can form by the core collapse of a dying star, but theories about the origins of supermassive black holes vary. Some think they were formed by mergers of many smaller stellar-sized black holes in the early days of their host galaxies. Others argue that the huge black holes formed first by some other process and galaxies formed around them.
    - It's worth noting that it might be easier to form a giant black hole than a smaller one. To form a black hole, a large mass is needed inside a small volume, inside its own Schwarzschild radius. From there, the laws of spacetime curvature take over.

    - To make a black hole of some mass M, a certain density of matter must be achieved. Suppose we're trying to make a black hole 10 times as massive, a mass of $10M$. That would have a Schwarzschild radius 10 times as large. Therefore, in some sense, within that radius, there is 1000 times the volume. To create a black hole of mass $10M$, the density we must have is only 1/100 the density we need to create a black hole of mass M.

Some have proposed that ultramicroscopic black holes might be created in the Large Hadron Collider at the CERN laboratory in Switzerland, but without a significant amplification of gravity, that simply could not happen.

- o For the same reasons, a tiny black hole could not form by the collapse of a massive star. To get a tiny mass within its own tiny event horizon, we would need to create quite incredible densities.

- Black holes are a real and fascinating part of modern astrophysics, but Einstein realized as early as 1917 that general relativity can be used to understand something even grander: cosmology, the scientific study of the universe as a whole.

Suggested Reading

Thorne, *Black Holes and Time Warps*, chapters 7–9. Thorne gives a great firsthand account of the rise of relativistic astrophysics, including the remarkable discoveries about black holes from the 1960s through the 1980s.

Tyson, *Death by Black Hole*, chapter 33. What happens to you when you fall into a black hole?

Wald, *Space, Time and Gravity*, chapters 6–9.

Wheeler, *A Journey into Gravity and Spacetime*, chapter 12. Wheeler, of course, was the man who introduced the term "black hole" in the first place.

Questions to Consider

1. Some gravitational physicists altogether eschew talking about what happens "inside" a black hole. Instead, they describe the black hole entirely in terms of the properties of its event horizon "surface" and the spacetime outside of it. The inside of a black hole, they say, is not really physical. From a purely philosophical point of view, do they have a point?

2. Suppose a miniature black hole of mass 10^{-6} kg formed inside a particle accelerator, and suppose this tiny black hole were initially at rest. Think about what would happen next. (If you like, you might try to estimate the size of the various effects of such a black hole. For example, how large is its event horizon compared to an atom?)

Problem

Calculate the radius of black holes with the mass of the Earth (6×10^{24} kg) and a galaxy of 10^{11} times the mass of the Sun.

Which Universe Is Ours?
Lecture 21

In this lecture and the next, we will apply Einstein's theory of gravity, general relativity, to the entire universe. After taking on supermassive black holes and active galactic nuclei, the universe itself is probably the only topic that could seem even more ambitious. The study of the universe as a whole is called cosmology. Modern scientific cosmology dates to 1917, less than two years after the Einstein equation, when Einstein himself began to explore the implications of his theory for the structure and dynamics of the cosmos.

A Large-Scale Picture of the Universe

- The distribution of matter and energy in the universe does not seem to be uniform. The stars have a great deal of matter in a comparatively small volume, while the space between the stars is almost empty. There are many stars within galaxies but very few in between. Galaxies themselves are found in clusters and even clusters of clusters. Even the distribution of dark matter is uneven.

- Although the universe is locally not very uniform, on the very largest distance scales, it is actually much the same everywhere. This fact is true as far out into space as we can observe, that is, more than 10 billion light-years.
  - This uniformity is sometimes called the cosmological principle: On very large scales, the universe, it is homogeneous (more or less the same at every point) and isotropic (more or less the same in every direction).

  - The big picture of the universe is fairly simple; all the individual objects of the universe—the planets, stars, and galaxies—are just details.

- According to Newton's law of gravity, what is the net gravitational force on our galaxy in this homogeneous and isotropic universe?

- This is a somewhat tricky question. After all, space has no end; there is an infinite amount of matter out there. On the other hand, there is about the same amount of matter in every direction.

- The most plausible answer to our question is that the net gravitational force on our galaxy is zero. It's not pulled this way or that way. The same should be true for any galaxy anywhere in the infinite universe.

- Of course, in our immediate vicinity, matter is clumped together; thus, there are local gravitational forces in one direction or another. The Earth is pulled toward the Sun as it orbits; the Sun is pulled toward the center of the galaxy as it moves; and so on. Galaxies move around within galaxy clusters, influencing each other and sometimes even colliding.

- In the really big picture, there are no net cosmological forces or accelerations. According to Newton's theory of gravitation, the universe might be roughly static. The details could be very dynamic, but the big picture might not change much over time. Such a universe could exist forever without a beginning or end.

Einstein's Picture of the Universe

- This Newtonian universe is the kind that Einstein had in mind in 1917 when he applied the Einstein equation to the universe.
  - Remember, in general relativity, gravitational forces are not the essential reality. If we adopt a freely falling frame of reference in the lab, those forces are zero.

  - The tidal effect, the relative acceleration of two freely falling objects, is important because the tidal effect is directly related to spacetime curvature.

- Einstein imagined a universe containing sparse uniform matter. The density is very low, but it is not zero. The universe is not empty. In words, the Einstein equation tells us that the total spacetime curvature is equal to a constant multiplied by the density of matter.

The constant involves G, Newton's gravitational constant. For the universe, because the density is not zero, the curvature is not zero. Therefore, parallel lines in spacetime must approach or diverge from one another. That's what curvature means.

- Imagine two galaxies in the universe some distance apart. Only gravity acts on the galaxies, so each galaxy follows a geodesic path in spacetime.
    - Because there must be some spacetime curvature, the relative acceleration of the two galaxies cannot be zero. In fact, according to the Einstein equation, that relative acceleration must be negative because the curvature is positive, like the density.

    - Thus, the galaxies might be moving closer together at increasing speed or moving apart at decreasing speed. In any case, the galaxies cannot stay put relative to each other.

- Each of the two galaxies is following a free-fall path, a straight line in spacetime. The relative acceleration of the two galaxies is a kind of cosmic tidal effect. Each galaxy is at rest with respect to the matter that is immediately around it. That means that if the distance between the two galaxies is changing, it's actually space itself that is expanding or contracting between them. Therefore, according to the Einstein equation, a homogeneous universe cannot be static. It must change over time.
    - The universe is either contracting or expanding. Einstein's theory of gravity predicts a changing cosmos. To Einstein, this conclusion seemed implausible. He thought that the cosmological prediction of his equation was a problem with his theory; thus, he adjusted the theory to solve the problem.

    - Einstein introduced a new constant, Λ, called the cosmological constant. In its new version, Einstein's equation tells us that total spacetime curvature plus the cosmological constant is equal to a constant multiplied by the density of matter.

- The constant Λ acts as a kind of cosmic antigravity. If it has just the right value, total spacetime curvature could be zero, at least on the largest scale, even if matter density is not zero. Then there would be no relative acceleration between two galaxies. Everything would balance, and the universe could remain static.

Conflicting Big Pictures

- Willem de Sitter, a Dutch astronomer, was fascinated by Einstein's general relativity and its cosmological implications, and he thought Einstein's original equation made more sense. His solution to Einstein's modified equation showed that a universe with a cosmological constant but little or no matter expands at an accelerating rate, driven by the cosmic antigravity of the Λ term.

- Alexander Friedmann, a Russian mathematician and physicist, developed his own mathematical solutions to the Einstein equation without Λ. These solutions describe different kinds of universes, all of which change over time.

- Georges Lemaître, a Belgian astronomer and physicist, considered mathematical model universes that were related to those of de Sitter and Friedmann and thought carefully about their physical meaning. Suppose the universe is expanding. If we extrapolate backward in time, that means that at a finite time in the past, space had zero size. All the galaxies were at the same point. This means that the universe itself has a finite age; there was a beginning to time. Lemaître had the idea of using the redshift of light emitted by galaxies to determine if the distance between galaxies is increasing.

- To recap, Einstein thought that the universe is probably static, and he was willing to modify his theory to allow this. De Sitter, Friedmann, and Lemaître each considered the possibility of a universe that expands, a more natural consequence of the relation between curvature and matter in the original Einstein equation.

The Cepheid Variable Yardstick

- In 1908, Henrietta Leavitt at Harvard University discovered a mathematical relation between period and luminosity for Cepheid variables—giant variable stars. Period and luminosity for these stars are connected by a simple rule: the longer the period, the brighter the star. This relation allowed scientists to determine the distance from such stars to Earth; in other words, it served as a yardstick for mapping galaxies, which would be a crucial step to applying Einstein's theory of gravity to the cosmos.

- In 1924, Edwin Hubble used the 100-inch-diameter telescope at the Mount Wilson Observatory to find Cepheid variables in M31, the great spiral in Andromeda. He measured their period and apparent brightness and proved that M31 is another galaxy far beyond the Milky Way. Throughout the 1920s, Hubble measured distances to other galaxies and their speed, using the Doppler effect. A remarkable pattern emerged.

- First, as Lemaître had noted, almost all galaxies are getting farther away; their light is redshifted. Second, roughly speaking, the farther away a galaxy is, the faster it moves away from us. In symbols, we write: $v = Hd$, where v is the speed of recession, d is the distance of the galaxy, and H is Hubble's constant.

- Although we've described Hubble's discovery as a kind of motion, what's really going on is that the space containing all the galaxies is getting larger. The universe is dynamic and changing; indeed, the very shape of spacetime is dynamic and changing.

The Progress of Gravitational Physics

- Our story about gravity seems like the most important story in all of science. Spacetime is curved. The cosmos itself is expanding. What could be bigger news than that? But in some respects, during the years after Einstein and Hubble, this branch of physics was a backwater. The main focus during the middle decades of the 20th century was on the physics of matter: quantum physics, nuclear

Before 1920, scientists did not know that Andromeda was a galaxy; some thought it might be a system of stars beyond the Milky Way—an island universe—and others thought it might be a nebula within the Milky Way.

physics, and particle physics. In gravitation and cosmology, progress was sometimes agonizingly slow.

- Even as recently as 30 years ago, disputes were ongoing about the basics of the expanding universe. How fast is the universe expanding? What is the value of Hubble's constant, that is, the relation between speed and distance in the expanding universe? The problem here is that to measure the overall expansion of the universe, we need to look at the big picture, and we reach a point at which we can't use the Cepheid variable yardstick because the stars are too distant for us to see, even with today's instruments.

- Gravitational physics began to blossom in the 1960s with new discoveries in astrophysics, particularly the discovery of the cosmic microwave background—electromagnetic radiation left over from the early universe. This microwave "noise" is a relic from a time when the universe was less than 0.5 million years old. Its discovery

began a new era of gravitational physics, allowing much more exact measurements and a much more detailed understanding of cosmology.

Suggested Reading

Dolling et al., eds., *The Tests of Time*, Part V ("Big Bang Cosmological Theory"). Contains many excellent readings by Leavitt, Hubble, de Sitter, Lemaitre, and others on the development of cosmology.

Wald, *Space, Time and Gravity*, chapters 4–5.

Questions to Consider

1. Consider a map of North America. If we consider squares 100 m on a side, the population of each square on the map greatly depends on where we draw the square. If the population in North America were distributed like the mass in the universe, then we should be able to find squares large enough so that the population is more or less the same no matter where we draw the square. Is the population of North America distributed in this way?

2. One objection to Einstein's "static" universe was that it was unstable. Without an exact balance between matter density and the cosmological constant, the cosmos would tend to expand or contract. Why are theories that require this sort of "fine-tuning" less satisfactory than other, more stable theories?

3. In this lecture, we described how gravitational physics (including cosmology) was in some respects a backwater of physics during the middle part of the 20^{th} century. What determines how important and active a branch of science is at any particular time? What changes and developments would promote a subject to greater importance?

Cosmic Antigravity—Inflation and Dark Energy

Lecture 22

As we saw in the last lecture, the universe is getting larger, and the things in it are getting farther apart. Looking back, we realize that the universe has a finite age, about 13.7 billion years by modern estimates. The very early universe was extremely hot and dense. The cosmic expansion is something like an explosion, except there is no edge to the explosion. It happens everywhere; space itself is getting larger. One early critic of this theory gave it the name big bang, which has stuck. This new theory explained the Hubble redshift law and the cosmic microwave background, but it also introduced puzzles of its own. In this lecture, we'll talk about some of the puzzles posed by big bang cosmology.

Origins and Composition of Galaxies

- Where did galaxies come from? Our earliest direct look at the universe is the cosmic microwave background, that is, light made up from a time when the universe was less than 0.5 million years old. The problem here is that the cosmic microwave background is extremely uniform, which means that in the early days, the matter in the universe was very evenly distributed. Now, however, matter has clumped together into galaxies and clusters of galaxies. How did that happen? The answer is gravity.
  - Recall from Lecture 11 the Jeans criterion for the collapse of a cloud of matter. It turns out that the early universe was not perfectly uniform. Some regions were slightly denser than others. The variations in density were small, but they were enough for gravity to work on. Not only could they collapse, but they could collapse fast enough to outrun cosmic expansion.

  - These regions contracted to form the first galaxies within 200 million years of the big bang, when the universe was less than 2% of its present age.

- Where did the chemical elements come from? The universe is mostly made up of hydrogen, the lightest element. The heavy elements, carbon, oxygen, iron, and so on, are formed inside stars by nuclear reactions. If we examine material that has never been in a star, we find a certain mix: 75% hydrogen and 25% helium, with traces of other elements. Evidently, that is the composition of the "stuff" out of which the galaxies formed.

- How did this original mix of elements come about? The answer is big bang nucleosynthesis.
  - In the first moments of the big bang, the universe was unimaginably dense and hot. Within a second or two, as the universe expanded and cooled to about 1 trillion degrees, protons and neutrons formed out of the energetic quark plasma that existed before. The protons and neutrons then underwent nuclear fusion reactions. Hydrogen was converted to helium and helium into heavier elements. These reactions went on only for a short time as the universe continued to expand and cool.

  - The speed of those reactions depended on the density of the protons and neutrons available, and the length of time those reactions lasted depended on gravity. More mass means that the expansion had to be faster in the earlier moments.

- The original mix of hydrogen and helium gives us a clue about the first few seconds of the universe, but when we work out the details, both the galaxy formation and the element formation puzzles reveal something interesting.
  - If we include only ordinary matter in our calculations—matter made out of protons, neutrons, and electrons—we cannot solve either puzzle.

  - There is not enough mass present to cause galaxies to form, and the early universe expands too slowly, so that more than 25% helium is produced. Thus, there must be another kind of matter present in the early universe: dark matter.

- The answers to our two puzzles tell us some important things about dark matter: It is not made of protons or neutrons; it must be cold, that is, its particles are not fast moving; and there must be about six times more dark matter than ordinary matter. At the moment, we do not know what dark matter is made of. It does not absorb or emit light, and it does not participate in ordinary nuclear reactions. We detect it only by its gravitational effects.

Uniformity and Flatness of the Universe

- As we've said, all large regions of the universe look the same. Why is this true? We might theorize that if different regions of the universe have interacted, then extra matter from one part might travel to another part, and things could even out. But the universe is not old enough for that to have happened.

- This puzzle is related to another one: Why is the space of the universe so flat? As you recall, Alexander Friedmann devised different mathematical universes based on Einstein's equation of spacetime curvature. In a Friedmann universe, the curvature of space depends on the total density of all kinds of matter and energy in the universe. When we observe the universe, it appears that its space is almost exactly flat over large scales. Why is that true?
    - In 1981, the American physicist Alan Guth proposed that the very early universe could have been quite different than it is today. Instead of something like a Friedmann universe filled with matter, it might have been more like a de Sitter universe, almost empty but with a cosmological constant. Einstein's Λ might have been significant for a tiny fraction of a second. In fact, there are some plausible ways for the physics of elementary particles to produce a huge value of Λ at least for a short time.

    - We called Einstein's cosmological constant, Λ, cosmic antigravity. In de Sitter's antigravity universe, the expansion happens at an accelerating rate. The size of the universe increases exponentially. Guth called this cosmic inflation, but

it didn't last long. It was all over when the age of the universe was only 10^{-30} s. But during that incredibly brief time, the universe expanded more than a trillion trillion times.

- Cosmic inflation means that widely separated points in space were once much closer to each other, which explains how the universe could be as uniform as we see it. Further, inflation would tend to flatten out any spatial curvature of the universe.

- The inflation idea is that the universe started out filled with a kind of energy. This energy acted as a cosmological constant, driving a very short period of very rapid inflation, but the energy was unstable and decayed into other kinds of energy. In fact, the energy released at the end of the cosmic inflation period could be the source of all the matter and energy that we see in the universe today.

- We have some evidence to support the theory of cosmic inflation. In essence, the end of inflation produced giant sound waves in the universe, and density variations in those sound waves affected the energy of light passing through the universe. This means that the photons of the cosmic microwave background record a little bit of information about the matter they pass through. Photons coming from slightly denser regions have slightly higher energies.
  - In fact, data from the Wilkinson Microwave Anisotropy Probe show that the cosmic background does have measurable density variations, and the characteristics of those variations match inflation theory closely.
  - Interestingly, those slight variations in density produced at the end of the inflation period are just enough to initiate the formation of galaxies in the early universe.

- The inflation theory is a bold hypothesis. It posits that matter, energy, and gravity were all once quite different. In the first 10^{-30} s, the universe was dominated by cosmic antigravity. That brief era and its aftermath have shaped everything we see.

The Shape of the Universe Today

- How has gravity affected the evolution of the universe? We answer this question by appealing to a basic principle of observational cosmology: When we look farther out into space, we are looking farther back in time.

- Imagine that we extend Edwin Hubble's diagram of redshift velocity versus distance for very distant galaxies. We include galaxies billions of light-years away. We do not expect to get a straight line because the speed of expansion has changed. The shape of the curve we get will tell us about the history of cosmic expansion.
    - The problem here is that we need a different yardstick to determine the distance to such distant galaxies. Cepheid variable stars are not bright enough to see more than 200 million light-years away.

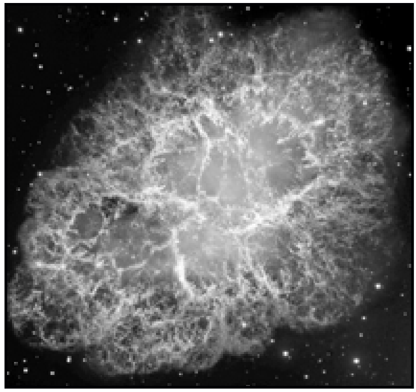
    - We need something that is brighter but with a known luminosity so that we can use its apparent brightness to determine its distance. Luckily, nature provides such objects: supernovas.

- Supernovas are gigantic explosions of stars. They come in different types and are recognizable from their spectra, the curve of their brightness over time, and so on. Type Ia supernovas are those formed when the mass of a white dwarf reaches the Chandrasekhar limit, and all Type Ia supernova explosions should have about the same luminosity.

- At their brightest, Type Ia supernovas are about 5 billion times brighter than the Sun, but they occur infrequently. When we find a Type Ia supernova, we measure its peak brightness and use that to determine the distance to its galaxy. We then measure the redshift from the galaxy and determine its recessional velocity. That adds another point to the Hubble diagram of velocity versus distance.

- In the late 1990s, using this methodology, astronomers discovered that the expansion of the universe is accelerating. Apparently, the cosmological constant is not zero after all. Not only was cosmic

antigravity present at the beginning of the universe during inflation, but there's a tiny amount left over. That leftover Λ is pushing galaxies apart today at an ever-increasing rate.


© Stocktrek Images/Thinkstock.

A Type Ia supernova is formed when a white dwarf reaches the Chandrasekhar limit; at that point, it collapses and either forms a neutron star or explodes with a huge release of energy.

- Cosmologists do not actually call Λ cosmic antigravity. From thinking about inflation theory and the energy release at the end of inflation, it seems more natural to think of Λ as a kind of energy density that is present somehow even in empty space. We do not see this energy density directly—only its effects on the expansion of the universe, its gravitational effects. It is dark, like dark matter, but to distinguish the two, we call it dark energy.
    - o Dark energy has some strange properties. As the universe expands, the density of ordinary matter (including dark matter) gets smaller because the same particles are spread out over a larger volume. The density of radiation energy—light, the microwave background, and so on—gets smaller even more rapidly as the universe expands because the radiation both spreads out in space and undergoes redshift.

    - o However, as the universe expands, the density of dark energy does not change. Dark energy is the energy of empty space. As space expands, there is more of it. Early in the universe, during nucleosynthesis, dark energy was an insignificant component of the cosmos. Several billion years ago, dark energy and matter had about the same density, but today, there is more than

twice as much dark energy as matter in our universe, and that's why the expansion is accelerating.

- What lies in the future of a universe that contains dark energy? As the universe continues to expand, the dark energy will, relatively speaking, become more important. The universe will look increasingly like de Sitter's universe, nearly empty of matter but with an accelerating expansion driven by dark energy.

- This model of the universe is called the Λ-CDM model; Λ stands for dark energy and CDM for cold dark matter. It is the best model we have for explaining the structure and evolution of the universe, but it does not explain the nature of either dark matter or dark energy.

Suggested Reading

Gates, *Einstein's Telescope*, chapters 7–10.

Guth and Steinhardt, "The Inflationary Universe," in *The Tests of Time* (Dolling et al., eds.). A popular account of the theory of cosmic inflation by two of its creators.

Questions to Consider

1. When we consider the production of hydrogen and helium in the early universe or the physics of the universe at the time when the cosmic microwave background formed, we do not need to consider the dark energy component of the universe. Why not?

2. We think that the cosmic antigravity in the universe remains constant over time. But some physicists have speculated that it might be increasing, growing stronger over time. Suppose that the density of dark energy is actually doubling every few billion years, even as the cosmos expands. What will the far future of our universe be like?

The Force of Creation

Lecture 23

In the last lecture, we answered some cosmological questions. For example, we saw how gravity amplified small variations in density in the matter of the early universe so that galaxies, stars, and planets could begin to form. Without that, life as we know it would have been impossible. In a real sense, then, gravity is the force that is ultimately responsible for life. In this lecture, we will take a deeper look at the properties of gravity that caused this to happen. We will see how gravity—the same force that pulls the apple down and steers the stars in their courses, that bends light and distorts time and space, that regulates the expansion of the universe—is also nature's ultimate force of creation.

A Unique Property of Gravity

- Let's begin by imagining a cluster of stars held together by mutual gravitation. In Newtonian terms, the cluster has both KE and PE. KE is positive and depends on how fast the stars are moving. PE is negative and depends on how closely the stars are clumped together. Both PE and KE might change a bit as the stars move around, but the total mechanical energy is constant. If the cloud is bound together by gravity, that total is negative.

- In Lecture 11, we learned about the virial theorem, which tells us that on average and in the long run, there is another relation between KE and PE, a quantity we call fugacity (f); $f = 2(\text{KE} + \text{PE})$ is about zero on average.

- What happens if we remove some energy from our cloud of stars? Reducing KE makes f negative, and the cloud contracts. As it contracts, gravity causes the stars to speed up; after a while, a new equilibrium is established. Once more, $f = 2(\text{KE} + \text{PE})$ is about zero.

- PE is more negative because the cloud is smaller; that means that KE is more positive. The stars are moving faster than before. That

seems rather strange. If we regard the cloud of stars as a kind of gas, KE would be the temperature of the gas. Faster particles mean a higher temperature. When we remove energy from the cloud, we are cooling it down by removing heat. But the temperature of the cloud is higher than before.

- This bizarre behavior relies on some unique features of gravity. In Newtonian terms, gravity is a universal, long-range, attractive force between masses, and no other force in nature has that combination of properties.
    - All life on Earth depends on light from the Sun. Our solar energy is generated by hydrogen fusion in the core of the Sun. Those fusion reactions require extremely high temperatures. But the original gas cloud forming the Sun and the planets was very cold, only 10° or 20° above absolute zero.

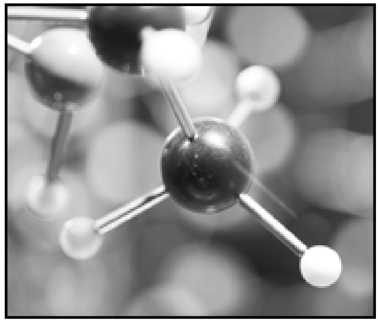
    - Where did the high temperatures come from to begin the nuclear reactions? The answer, of course, is gravity. The original gas cloud was larger than the Jeans mass for its density and temperature; thus, rather than dispersing, the cloud began to contract. The gas increased in speed and heated up. Much of that energy was radiated into space, but that made the cloud contract and heat up even more. Eventually, the center of the protostar in the middle of the cloud grew hot enough to start nuclear reactions.

    - Gravity started with a sparse, cold cloud of gas, concentrating it and heating it up by huge amounts, until a star was born, the Sun.

Gravity and Entropy

- The first law of thermodynamics is that energy is conserved. The second law is sometimes phrased this way: Entropy cannot decrease; it can only stay the same or increase. One way to think of entropy is as a measure of an object's microscopic disorder. Another way of defining entropy is as a measure of information, specifically, how much information we lack about the molecular-level details of an object.

- For example, we know only a few things about the air in a room: its total mass, chemical composition, pressure, temperature, and humidity. We do not know the exact positions and velocities of all the molecules in the air. That missing information means that the system has high entropy.



The positions and velocities of air molecules in a room are completely disordered; in other words, the air in a room has high entropy.

- The second law of thermodynamics says that entropy tends to increase, not decrease. If the air in the room starts out all in one corner, a situation of lower entropy, it will tend to spread out into the whole room, a situation of higher entropy. Higher entropy does not spontaneously lead to lower entropy; if the air starts out spread throughout the room, it does not spontaneously gather together in one corner.

- The air in a room also has a fairly uniform temperature; that means that thermal energy is distributed more or less evenly among the molecules in different parts of the room. It isn't that the molecules all have exactly the same energy. The gas contains both slow molecules and fast molecules, but molecules in different parts of the room are equally likely to have a given amount of energy.

- The entropy of a gas will be greatest if it is spread out evenly throughout its volume and at a uniform temperature; according to the second law of thermodynamics, any cloud of gas should tend toward this situation. But that raises a perplexing question about cosmology.

- How did we get the variety and complexity we see today out of a smooth, uniform early universe? The answer is gravity. As the cosmos expanded and cooled, the expansion was controlled by gravity, acting at a cosmic scale. Gravity also caused local parts of

the universe to contract and heat up. Then, galaxies formed and, within galaxies, stars and all the rest.

  - The law of entropy does not absolutely prevent the creation of order. It simply requires that a price must be paid for order.

  - Entropy can decrease in one place as long as it increases someplace else. There is no place outside the universe for us to put extra entropy.

- The universe seems more ordered now than it did billions of years ago when the cosmic microwave background formed. The driving force creating that order has been gravity. Does this process contradict the second law of thermodynamics? Does gravity make entropy decrease? The answer is no.

- Entropy works a little differently for systems when gravity is important.
  - Consider again our collapsing cloud of stars. The stars at the end occupied a smaller volume, which would tend to reduce entropy. Because of gravity, the stars also have a higher speed on average, which would tend to increase entropy.

  - Because of gravity, smaller and hotter can sometimes mean a higher entropy than larger and cooler. In other words, gravity can change the entropy equation so that a less uniform distribution of matter and energy is actually favored by the second law of thermodynamics. That's how galaxies, stars, and planets could form in the first place.

Testing Thermodynamics with a Black Hole

- Consider an ordinary object, such as an apple. The mass of the apple represents a very large energy content, according to Einstein's relation between mass and energy: $E = mc^2$. The apple also has entropy. According to the laws of thermodynamics, energy is conserved and entropy cannot decrease.

- Now we toss the apple into a black hole. It falls through the event horizon. Tidal forces tear it to pieces. The pieces are crushed to zero size in the singularity. The apple is gone. What happened to the energy and entropy of the apple?

- The energy of the apple is still present in the mass of the black hole. Even though the matter that went into a black hole has disappeared, the gravitational field around the black hole remains. The strength of that field tells us the mass of the black hole. When the apple falls into the black hole, the gravitational field increases. The mass of the black hole is greater, and because mass is a measure of energy content, the black hole's energy has increased. The first law of thermodynamics still holds.

- What about the second law of thermodynamics? Hasn't the entropy of the apple disappeared? A black hole is an incredibly simple object, much simpler than an apple. It is nothing but a naked gravitational field left over from a completely collapsed object. All the details about the matter that formed the black hole are completely gone from the universe. Without any microscopic details, there is no entropy. The entropy of the apple has vanished. The second law of thermodynamics is in trouble.

- This was the dilemma that an Israeli physicist named Jacob Bekenstein pondered in the early 1970s. He had the idea that we should calculate the entropy of a black hole slightly differently than we calculate the entropy of an apple. For Bekenstein, the entropy of a black hole is the amount of information we lack about the matter and energy that crossed over the event horizon as the black hole formed. In other words, black hole entropy is the information that has been lost in the black hole.

- Using quantum physics, Bekenstein found that the entropy of a black hole is proportional to the area of its event horizon.
  - This is a surprising connection between two different kinds of ideas. The event horizon is just a mathematically defined surface of empty space, but it's a special surface. Information

can cross it in only one direction, inward. That's why the horizon area is related to entropy. The entropy of the black hole is just the total amount of information that has crossed that surface since the black hole formed. Bekenstein's rough calculations showed that this idea works out.

- The entropy of a black hole is a staggeringly large number, immensely greater than the entropy of the Sun. When a star collapses to form a black hole, its entropy increases dramatically, which is, of course, perfectly consistent with the second law of thermodynamics.

- Finally, Bekenstein's idea seems to agree with an earlier discovery about black holes made by Stephen Hawking: the idea that the total black hole horizon area never decreases.

Hawking versus Bekenstein

- Hawking initially believed that Bekenstein was wrong. If something has both energy and entropy, it must also have temperature. If it has a nonzero temperature, then heat can flow from it to something colder. But a black hole can never transfer heat to anything else. In effect, Hawking argued, a black hole's temperature must be exactly absolute zero. For this reason, Hawking concluded that a black hole cannot have a real entropy.

- In analyzing the quantum physics of matter and energy in the close vicinity of a black hole, Hawking discovered that black holes are not quite black. There is a very faint outward radiation from a black hole, exactly as if it had a real temperature. The larger the black hole, the lower the Hawking temperature.

- A black hole the mass of the Sun or larger would certainly not be in equilibrium in our universe. It would absorb more energy from the cosmic microwave background than it would radiate outward. The black hole would slowly grow in size.

- Hawking's discovery confirmed Bekenstein's original hypothesis. Black holes do have entropy and temperature, and the entropy of a black hole is proportional to the surface area of its event horizon. When we combine quantum physics and curved spacetime, we find many new insights about gravity, information, and entropy.
    - Because of gravity, a smooth and uniform early universe can develop into the clumpy and varied universe we live in. Because of gravity, increasing clumpiness can also increase entropy.
    - That fact about gravity is the ultimate reason for the order and complexity around us. A black hole is just the most extreme version of the same fact. Matter in the black hole is so clumped together that it is squeezed into a point singularity. Its entropy still exists in the area of its event horizon.

Suggested Reading

Thorne, *Black Holes and Time Warps*, chapter 12.

Wald, *Space, Time and Gravity*, chapter 10.

Questions to Consider

1. Cosmologists sometimes refer to the "dark ages" of the cosmos, the long period of the history of the universe between the creation of the cosmic microwave background (less than 0.5 million years after the big bang) and the formation of the first stars and galaxies (almost 0.5 billion years later). Was anything important happening during this time?

2. Before the discovery of dark energy, it seemed possible that the universe would expand, halt, and then recollapse after a few hundred billion years. Given that entropy would continue to increase all during that cycle, how would the universe look different during its collapse (the big crunch) than it did after the big bang?

The Next Revolution

Lecture 24

So far in the course, we've seen two revolutions in our understanding of gravity. First, Newton showed that gravity is an attractive force that acts between every mass and every other mass. Then, Albert Einstein showed that gravity is really an effect of curved spacetime. In the first decades of the 20^{th} century, the theory of quantum mechanics was developed, which describes the behavior of elementary particles. Quantum mechanics is still the most successful and widely applicable theory of physics, but we do not yet have a good theory of quantum gravity. What we need is another revolution.

Reconciling Quantum Mechanics and Gravity

- If quantum mechanics is to be logically consistent, it must apply to everything.
    - For instance, in quantum mechanics, the location of a particle can be indeterminate. In some situations it is impossible, even in principle, to tell whether an atom or an electron is in one place or another, but even an atom or an electron has a gravitational field.

    - The presence of the particle affects spacetime around it in a tiny way; thus, either the particle has a definite location, which would mean that there's something wrong with quantum mechanics, or the curvature of spacetime is also somewhat indeterminate, which would mean that there's something wrong with Einstein's general relativity.

- When Steven Hawking used quantum theory to show that black holes are not quite black, he was actually applying quantum theory to everything but gravity. In other words, we know something about quantum mechanics in warped spacetime but not much. We need a quantum description of spacetime itself.

- We have some clues about a quantum description of spacetime. For instance, there are three fundamental constants of nature: the speed of light, *c*, relating time and space, energy and mass; the gravitational constant, *G*, connecting mass with spacetime curvature; and Planck's constant, *h*, which connects photon energy and frequency.
  - Planck's constant sets the physical scale of the quantum world. Because this constant has a very tiny value, the strange effects of quantum mechanics are typically very small compared to our everyday experience.
  - Planck himself noted that these three constants, *c*, *G*, and *h*, can be put together into a length, now called the Planck length; this tiny length sets the basic scale of nature: $L_p = \sqrt{\dfrac{hG}{c^3}} = 1.6 \times 10^{-35}$ m, which works out to be about 4×10^{-35} m.
  - If we could examine the world at that kind of tiny scale, the simple space and time we know would turn out to be something quite different. To understand physics at length scales that small, we need to understand quantum gravity.

- Another clue about quantum gravity comes from the application of quantum physics to light.
  - Electromagnetic waves carry their energy in the form of particles called photons. The electromagnetic field is somehow both a smooth, continuous wave and discrete packets of energy; this double nature is basic to the quantum physics of light.
  - Physicists conjecture that the same thing must be true for gravity. In Lecture 19, we discussed gravitational waves, moving ripples in the curvature of spacetime. Physicists think that there must be a kind of particle involved called a graviton.

- In the search for quantum gravity we have two special problems: We lack experimental data, and we aren't quite sure what guiding principles to follow. It may be that some aspect of gravity we

have discussed in this course will turn out to be the key to solving the puzzle. Physicists are now pursuing many different lines of attack on the problem, including string theory, M-theory, and loop quantum gravity.

A Return to Entropy

© iStockphoto/Thinkstock.

The energy added when stretching a rubber band goes into increasing the entropy of molecular motion to offset the decreased entropy of the straightened polymers.

- Think about stretching a rubber band; that action exerts a force on your fingers. In other words, you need to do work to stretch it; you must put energy into the rubber band. What is it about the rubber band that makes this necessary?
    - You might think that when you stretch the rubber band, you move the rubber molecules out of place, and because there are molecular forces holding the molecules in place, energy is required to shift them.
    - This explanation sounds plausible. Indeed, it's the correct explanation for why energy is required to stretch a metal spring, but it's wrong for a rubber band.

- A physical chemist named Eugene Guth worked out the real answer to this question. In the 1930s, Guth became interested in the thermodynamics of polymers, which are extremely long chains of repeated molecular units. Rubber is made out of polymer molecules that are mostly bundled up in complicated tangles. When you stretch a rubber band, you're actually straightening out the polymers a bit.

- By itself, straightening a polymer doesn't require any energy, but a straight polymer is more ordered than a tangled one. If the polymer is straightened out, we know more about where the pieces are. The real difference here is entropy. A tangled polymer has higher entropy, and a straight polymer has lower entropy.

- But entropy cannot just decrease. If the entropy of the polymer configuration goes down, some other sort of entropy must go up. In the case of the polymers in the rubber band, the molecules must start to vibrate faster. To stretch the rubber band, therefore, we must heat it up, and that is what demands an energy input.

- The energy you add to the rubber band does not go into straightening the polymers. It goes into heating up the rubber—raising the entropy of molecular motion to offset the lowering of the entropy of the polymer tangle. The polymer molecules, as it were, pull themselves toward the higher-entropy situation. In effect, Guth showed that entropy all by itself can lead to a force.

- Quite recently, a Dutch physicist, Erik Verlinde, has combined Guth's idea of an entropic force with the holographic principle of another physicist, Gerard 't Hooft.
  - Consider any region of space bounded by a surface, for example, a room. The quantum state of the contents of the room constitutes a certain amount of information. From one perspective, this information is located inside the room, but 't Hooft said that there was another possibility.

  - As far as the rest of the universe is concerned, that information could just as easily be located on the boundary of the room, that is, the surface formed by the walls, floor, and ceiling, in the same way that information in a black hole is associated with its bounding surface. The three-dimensional room itself is a kind of quantum hologram cast by the two-dimensional boundary. The same thing is true for any volume of space.

- In the last lecture, we saw that gravity makes entropy favor things contracting rather than expanding. According to Verlinde, this may be what gravity actually is: an expression of the second law of thermodynamics.
  - The entropy of the universe is higher when things are closely clumped together, and the second law tells us that this is how things will tend to go. In this view, gravity is an entropic force, an expression of this tendency.

  - If Verlinde is right, then the universal force of Newton and the curved spacetime of Einstein might turn out to be expressions of entropy, consequences of the second law of thermodynamics.

Summing Up Our Study

- Our study of gravity has taken us on some strange journeys since we first began. Consider what we now know about the simple fall of an apple. It falls with a constant acceleration, about 10 m/s^2, near the Earth. In fact, every falling object near the Earth accelerates in the same way.

- Newton told us that the Earth exerts a gravitational force on the apple, exactly the same kind of force that steers the Moon in its orbit around the Earth. If the apple were farther from the Earth, the strength of the force of gravity would be less. In fact, if it were twice as far from the center of the Earth, the force would be four times less.

- The force of gravity exerted by the Earth is directed toward the center of the Earth because every part of the Earth, near and far, pulls on the apple. Because the Earth is roughly spherical, the net effect is to pull the apple straight down toward the center.

- Newton had two distinct concepts for mass for the apple: inertial mass, which measures how much the apple resists acceleration by any force, and gravitational mass, which measures how strongly the apple is affected by gravity. Nevertheless, for the apple, and for

everything else, these two masses are exactly equal, a fact called the principle of equivalence.

- As the apple falls, its total mechanical energy is conserved. Its KE increases as it speeds up, and its PE decreases as the apple and Earth get closer together; KE + PE stays constant. If we have two apples falling at different locations, their accelerations will be slightly different; they'll be slightly pulled apart or together. This is the tidal effect, the same effect responsible for the rise and fall of the oceans.

- Einstein had a different way of thinking about the fall of the apple. He said that if you imagine yourself falling next to the apple, close enough to ignore the tidal effect between you, you would not see the apple accelerate at all. A freely falling frame of reference has zero gravity.

- In a similar way, if you were far out in deep space in an accelerating laboratory, then the apple would appear to fall in the opposite direction from the acceleration. That's an effect of inertia—the apple tends to lag behind the accelerating lab—but it is indistinguishable from gravity. That's Einstein's version of the principle of equivalence.

- In fact, Einstein's view is that the apple in free fall follows a straight world line, a geodesic path in the curved spacetime near the Earth. That path is a path of greatest clock time. If you strap a watch to the apple and consider all the ways the apple might go, from its initial where-and-when to its final where-and-when, that free-fall path is the path that would maximize the time registered by the watch. This works because of gravitational time dilation, the fact that lower clocks run more slowly.

- The resulting world line of the apple seems curved because of our spacetime map projection. Spacetime curvature is actually measured by the tendency of parallel world lines to converge or diverge, that is, by the tidal effect. In empty space, the total

curvature over all three directions is zero, but when matter is present, curvature is governed by the Einstein equation. Total spacetime curvature is a constant multiplied by matter density. In Wheeler's motto: Spacetime tells matter how to move; matter tells spacetime how to curve.

- The inertial frame of the falling apple is slightly dragged along with the Earth's rotation due to the Lense-Thirring effect, and from the point of view of the apple, the whole universe appears to be turning very slowly. Only in the last few years has this extremely tiny effect been measured near the Earth by Gravity Probe B.

- If we drop the apple into a black hole, we will never see it cross the event horizon, but if we follow along the apple's world line, it does cross the horizon and, soon afterward, is crushed to zero volume or to the quantum-gravity scale at the central singularity. Finally, it might turn out that the gravitational effect on the apple, the curvature of spacetime, that causes the apple to fall is really due to a quantum entropic force analogous to the elastic forces of a rubber band. We've seen all of that in the fall of an apple.

Suggested Reading

Feynman, *The Character of Physical Law*, chapter 7. Though not speaking directly about the problem of quantum gravity, Feynman makes some wise and insightful observations about how new science is discovered.

Smolin, *Three Roads to Quantum Gravity*.

Verlinde, "On the Origin of Gravity and the Laws of Newton."

Questions to Consider

1. At the beginning of our course, we said that the gravitational force had four essential characteristics: (1) Gravity is long range. (2) Gravity is very weak. (3) Gravity is attractive. (4) Gravity is a universal force affecting all masses. How do these observations relate to the revolutionary discoveries of Galileo, Newton, and Einstein?

2. Here's a fun demonstration that you can do at home: Hang a weight from a rubber band. Now heat the rubber band—add entropy to it—using an electric hair dryer. Do you expect the stretched rubber band to get longer or shorter? Make a prediction based on our discussion of polymer entropy and then perform the experiment to check your prediction. (It helps to mark the original position of the hanging weight in some way to allow you to observe small changes in the length of the rubber band.)

3. In this lecture, we suggested that the odd relationship between gravity and entropy might be a clue to the next revolution in gravitational physics. Can you think of other properties of gravity that might suggest other possibilities?

Problem

If an atomic nucleus (radius = 10^{-15} m) were magnified to the size of a planet (radius = 10^{7} m), would the Planck length be larger or smaller than the original nucleus?

Abbreviations

arcsecond ....arcsec

astronomical unit ....AU

atmosphere..atm

centimeter ....cm

joule ....J

kelvin ....K

kilogram..kg

kilometer...km

liter...L

meter ....m

microsecond..μs

millisecond ....ms

nanosecond ....ns

newton ....N

pascal ....Pa

second ....s

Physical Constants, Lengths, Masses, and Speeds

Physical Constants

Speed of light: $c = 3.00 \times 10^8$ m/s

Newton's gravitational constant: $G = 6 \times 10^{-11}$ m^3/kg s^2

Coulomb's electric force constant: $k = 9.0 \times 10^9$ N m^2/C^2 (C = coulombs)

Planck's quantum constant: $h = 6.63 \times 10^{-34}$ J s
(The following is often used: $\hbar = \dfrac{h}{2\pi} = 1.05 \times 10^{-34}$ J s)

Lengths

Radius of the Earth: 6400 km = 6.4×10^6 m

Earth-Sun distance (AU, or astronomical unit):

150 million km = 1.5 × 1011 m

1 light-year: 9.5 trillion km = 9.5×10^{15} m

1 parsec: 3.26 light-years = 3.1×10^{16} m

Schwarzschild radius of a 1-solar-mass black hole: $R_S = 2.95$ km

Masses

Mass of Earth: 6.0×10^{24} kg

Mass of Sun: 2.0×10^{30} kg

Speeds and Accelerations

Escape speed from Earth: 11.2 km/s

Circular orbital speed near Earth: 7.9 km/s

Acceleration of gravity near Earth: $g = 10$ m/s^2

Equations of Gravity

Falling Bodies

Velocity: $v = \Delta x/\Delta t$

Acceleration: $a = \Delta v/\Delta t$

Galileo's law for distance covered during constant acceleration:
$x = \frac{1}{2}at^2$ (x = distance fallen, a = acceleration, t = time)

Acceleration due to gravity near Earth: $g = a = \frac{GM}{R^2}$
(a = acceleration, G = Newton's constant, M = mass of the Earth, R = radius of the Earth)

Relative acceleration: $a = \frac{GM}{r^2}\left(\frac{2h}{R}\right)$
(M = mass of the Earth, h = distance between two falling bodies, r = average distance from Earth, R = radius of the Earth)

Orbits

Kepler's harmonic law: $a^3 = KP^2$
(a = semimajor axis, P = orbital period, K = constant)

Orbital period: $P = \sqrt{a^3 / K}$
(P = orbital period, a = semimajor axis, K = constant)

Newton's harmonic law: $a^3 = \left(\frac{G}{4\pi^2}\right)MP^2$
(a = semimajor axis, G = Newton's constant, M = mass, P = orbital period)

Angular momentum: $\ell = mrv$
(ℓ = angular momentum, m = mass, r = radius, v = speed)

Circular orbital speed: $\sqrt{\frac{GM}{r}}$
(G = Newton's constant, M = mass, r = radius)

Newton's Three Laws of Motion

1. **Law of inertia (from Galileo):** In the absence of an external force, a body maintains a constant velocity.

2. **Law of motion:** $F = ma$ [net force on a body = (mass of the body) × (acceleration of the body)]

3. **Law of action and reaction:** If object A exerts a force on object B, then B exerts a force on A that is equal in strength but opposite in direction.

Universal Attraction

Gravitational attraction between two particles: $\dfrac{Gm_1m_2}{r^2}$

(G = Newton's constant, m = masses, r = distance apart)

Force of gravity on Earth: $F = \dfrac{GMm}{R^2}$

(F = force, G = Newton's constant, M = mass of the Earth, m = mass of an object, R = radius of the Earth)

Acceleration of gravity on Earth: $g = \dfrac{GM}{R^2}$

(g = acceleration, G = Newton's constant, M = mass of the Earth, R = radius of the Earth)

Escape speed: $\sqrt{\dfrac{2GM}{r}}$ (G = Newton's constant, M = mass, r = radius)

(Relative acceleration due to) tidal effect: $a = \dfrac{GM}{r^2}\left(\dfrac{2h}{r}\right)$

(a = relative acceleration, G = Newton's constant, M = mass, h = separation between two bodies, r = radius)

Coulomb's law for force between two electrical charges: $F = \frac{kQ_1Q_2}{R^2}$

(F = force, k = constant expressing the strength of the electric force, Q = charges [in Coulombs], R = distance apart)

Energy

Kinetic energy: KE = $\frac{1}{2}mv^2$ (m = mass, v = speed)

Potential energy: PE = $-\frac{Gm_1m_2}{r}$ (G = Newton's constant, m = mass, r = radius)

Fugacity: f = 2 KE + PE

Gravity as Spacetime

Energy: $E = mc^2$ [energy = mass × (speed of light)^2]

Photon energy: $E = hf$
(E = photon energy, f = light frequency, h = Planck's constant)

Einstein's equation (simplified form): total spacetime curvature =

$\mathcal{R}_{\text{total}} = \left(\frac{4\pi G}{c^2}\right)\rho$ (where ρ = density of matter)

Spacetime interval: $s^2 = c\Delta t^2 - \Delta x^2$
(s = spacetime interval, c = speed of light, t = time)

Deviation equation: $a = -\mathcal{R}v^2d$

(d = distance between two particles, v = speed along initially parallel lines, a = relative acceleration [positive as the particles tend to diverge and negative if they converge], $\mathcal{R}$ represents the curvature [positive if a is negative and negative if a is positive])

Event horizon radius: $R_s = \dfrac{2GM}{c^2}$

(R_s = Schwarzschild radius, G = Newton's constant, M = mass of black hole, c = speed of light)

Galactic expansion: $v = Hd$

(v = speed of recession, d = distance of the galaxy, and H = Hubble's constant ≈ somewhere between 21.5 to 23.4 km/s per 1,000,000 light-years)

Planck length: $L_p = \sqrt{\dfrac{hG}{c^3}} = 1.6 \times 10^{-35}\,\text{m}$

(L_p = Planck length, h = Planck's constant, G = Newton's constant, c = speed of light)

Answer Key—Questions to Consider

Lecture 1

1. Suppose human civilization moves away from the surface of the Earth and establishes itself in deep space, where gravitational forces are negligible. How would the lack of gravity affect the activities of daily life, agriculture, architecture, the arts, and athletic competitions—to mention only the A's?

Answer: This open-ended question is intended to inspire reflection; answers may vary. My own reflections involved comic scenes of varying activities:

- When I'm boiling water in the kitchen, how do I keep the pot on the stove and the water in the pot? Once the water next to the sides of the pot boils, the expanding steam will push the rest of the water out!

- Suppose I wish to plant a garden. My plants require dirt on one side (for the roots) and light and air on the other. What holds the dirt in place? Many plants can be attached to a trellis to guide their growth. But how do I water them?

- The rules of baseball presume gravity at every point. Gravity keeps the players on the ground; players push against the ground when throwing, hitting, and running. Without gravity, pretty much every hit is a home run!

On the other hand, a lack of gravity might provide some remarkable opportunities:

- Architecture, freed from the structural constraints of supporting huge weights, could employ new materials in buildings, which could assume new, delicate forms.

- In the arts, a lack of gravity might allow amazing new types of dance. The science fiction story "Stardance" by Spider and Jeanne Robinson envisions just such a "zero-gravity" ballet.

(The story won the Hugo award for the best science fiction novella of 1978.)

2. Some physicists have speculated about the existence of a "fifth force" that would be nearly as weak as gravity but act only over a short range. Such a force would be extremely difficult to detect. Why? (As of this writing, no experimental search has found such a force.)

Answer: Gravity is a weak force, but because it is long range, we can easily observe the gravitational force exerted by the entire Earth. If a fifth force is both weak and short range, only objects very close by would contribute to the observed effects.

Lecture 2

1. What is the logical connection between the principles of inertia and relativity, both discovered by Galileo? To illuminate this question, suppose there were some object for which the law of inertia did not hold, that is, an object that was "naturally at rest," as Aristotle supposed. How would such an object behave in Galileo's shipboard laboratory?

Answer: An "Aristotelian" object in a moving laboratory would appear to move toward the back of the laboratory, as Aristotle supposed. Thus, it would be possible to distinguish whether the shipboard laboratory was moving or sitting still, contrary to the principle of relativity.

2. Consider two objects, A and B. A falls twice as far as B, so it hits the ground at a higher speed. Some natural philosophers before Galileo believed that the object falling twice as far would be moving with twice the speed when it hit. Is this true?

Answer: No. The speed of a falling object increases with time according to the equation $v = at$. If A hits twice as fast as B, it must fall for twice as much time. But the distance fallen is $x = \frac{1}{2}at^2$; thus, an object that falls for twice as much time falls four times as far. If A falls twice as far as B, it hits the ground at only about 40% faster than B.

Lecture 3

1. Use the astronomer's trick mentioned in the lecture—that the width of a little finger at arm's length is about 1°—to consider the accuracy of Tycho's astronomical measurements. His instruments could locate a star or planet to within 1/30°. Can you resolve an angle that small with your own unaided eye?

   Answer: Just barely. One way to measure the resolving power of the eye is to ask how far apart two stars must be for us to see them as two separate objects rather than one. For normal human eyesight, this separation is about 1 minute of arc, or 1/60°.

2. It is sometimes said that Kepler's third law of planetary motion, the harmonic law, "completed" his theory. Explain this claim. What ingredient was missing in the first two laws, and how did the third law supply it?

   Answer: The law of ellipses gives the shape of each orbit, and the law of equal areas tells how the motion of each planet varies as it moves around its orbit. But these laws only refer to each orbit individually; they do not tell how the motion of one planet is related to the motion of another in a different orbit. By connecting orbital period to semimajor axis, the harmonic law related the orbital speeds of different planets to one another.

Lecture 4

1. Newton's law of universal gravitation—that every mass exerts an attractive force on every other mass—was an amazing leap of imagination. What reasons could Newton have provided for adopting such a radical hypothesis?

   Answer: One clue would be Galileo's discovery of the law of free fall. Because every object falls with the same acceleration, each km of mass is pulled with the same force by the Earth's gravity. By Newton's third law of motion, therefore, each km of mass must exert the same force on

the Earth. Because all bodies experience gravitational force, all bodies must also exert it.

2. Two planets have exactly the same total mass, but one is denser than the other one and, thus, has a smaller radius. Which one has a greater surface gravity?

Answer: The denser one has the greater surface gravity; M is the same, but R is smaller, so that $F = \frac{GM}{R^2}$ is larger.

Lecture 5

1. Before Cavendish's experiment, the value of the gravitational constant G was not known. Newton nevertheless knew that the Sun is more than 300,000 times more massive than the Earth. How could he know this fact? Would his reasoning have relied on the measurement of solar system distances by Cassini and Richer?

Answer: By comparing the gravitational accelerations produced by the Earth (on the apple, say) and the Sun (on the orbiting Earth), Newton could compare the values of GM ($G \times M$) for the Earth and Sun. Because the gravitational constant G is the same for each, this gives the ratio of the masses. This, of course, depends on knowing the distance between the Earth and the Sun and, thus, depends on the measurements of Cassini and Richer.

2. Eotvos's experiments with the torsion balance confirmed the principle of equivalence to fantastic precision. Could such precision be obtained by observing the free fall of dropped objects, à la Galileo? Why or why not?

Answer: There are many reasons why a simple free-fall experiment would not be as precise as the torsion balance experiment. For one thing, it would be extremely difficult to measure free-fall time precisely enough. More importantly, the smallest variations in air resistance would affect the free-fall motion by much more than 1 part in 100 million.

Lecture 6

1. The next time you visit a fair or an amusement park, consider how energy, momentum, and angular momentum all play roles in various rides. The roller coaster is especially interesting, as are the various pendulum swing rides that have names ranging from Pirate Ship to Frisbee.

   Answer: One especially obvious example is the exchange between PE and KE in a roller coaster. At the top of a hill, the coaster is moving slowly, but when it rolls down the sloping track, it turns PE into KE and speeds up. Traditional roller coasters get started by pulling the cars with a chain to the top of the first—and highest—hill. Some modern coasters get started horizontally, with the mechanical energy added by an electric linear induction motor.

   Rotating rides also provide some excellent illustrations. In one type, riders move at the ends of arms mounted to a large vertical shaft that spins around. The arms move in and out, decreasing and increasing the radius of the circular motion. The speed is much faster when the arms are close in because the angular momentum in the ride is roughly conserved.

2. Is the total mechanical energy (KE + PE) of the solar system positive or negative?

   Answer: The total mechanical energy for the solar system is negative. This is why the planets remain bound to the Sun. (If the total energy were positive, the planets would be moving fast enough to escape the Sun's gravitational pull.)

3. We saw that the planet Saturn is approximately 100 times as massive as the Earth and 10 times the radius. This means that the force of gravity at Saturn's surface is about the same as on Earth. Is the escape speed from Saturn greater than, about equal to, or less than the escape speed from Earth?

Answer: The escape speed is $\sqrt{\frac{2GM}{R}}$. Thus, if M is 100 times larger and R is 10 times larger, the escape speed will also be larger (by a factor of $\sqrt{10} \oplus 3.2$.

Lecture 7

1. Consider Newton's cannon thought experiment. If the cannonball is fired fast enough, it goes in a circular orbit around the Earth (ignoring air resistance). What happens if the cannonball is fired a little faster than this? What happens if it is fired twice as fast?

Answer: If the cannonball is moving a little faster than the circular orbit speed, it follows an elliptical orbit around the Earth. The firing position is the point of nearest approach to the planet. If the cannonball is fired twice as fast, it will be moving faster than the escape speed and will follow a curved hyperbolic path into deep space.

2. Here are three types of spacecraft missions from Earth to Mars: To fly swiftly past Mars, to enter into orbit around Mars, and to land gently on the surface of Mars. Which one requires the most fuel? Which one requires the least? Explain why.

Answer: Each mission involves sending a spacecraft to Mars. But when the spacecraft approaches Mars, the gravity of the planet accelerates it. Without any slowing down, the spacecraft is actually moving faster than the escape speed of Mars. Thus, it requires the least amount of fuel to fly past Mars, and the most fuel to land softly on the surface.

Lecture 8

1. Planet A has a single moon. Planet B has a moon that is twice as massive but orbits twice as far away. Which planet has greater tides?

Answer: The tidal effect of a moon of mass M at a distance r is proportional to $\frac{M}{r^3}$. Thus, Planet B experiences a tidal effect that is

$\frac{2}{2^3} = \frac{1}{4}$ times as large as Planet A. We predict that Planet A has greater tides.

2. By Newton's third law, the gravitational pull of the Earth on the Moon is equal to the gravitational pull of the Moon on the Earth (though they are opposite in direction). Is the same equality true for the tidal effect of the Moon on the Earth and the Earth on the Moon?

Answer: No. The Earth-Moon distance is the same in each case, but because the Earth is much more massive, the tidal effect at the Moon due to the Earth is greater than the effect at the Earth due to the Moon. Although the Moon has slowed down Earth's rotation a bit over time, the Earth has slowed the Moon's rotation much more, so that the Moon is now "locked" to the Earth, keeping the same face toward it at all times.

Lecture 9

1. Anomalies in the orbit of Uranus were found to be due to the gravitation of a new planet, Neptune. Anomalies in the orbit of Mercury, however, were not due to a new planet. In one case, the puzzle could be resolved within the existing (Newtonian) theory of gravity, but in the other case, it could not. This brings up one of the most important questions in science: How can we tell whether a piece of anomalous data requires us to change our basic theory?

Answer: This is a profound question that has no cut-and-dried answer. In the cases of Uranus and Mercury, it was reasonable to guess that an unknown planet existed. Such a planet was quickly discovered in the case of Uranus but not in the case of Mercury. Therefore, another idea was required to solve the puzzle.

2. The first exoplanets discovered by the radial velocity method were so-called "hot Jupiters"—very massive planets extremely close to their

parent stars. However, such planets are believed to be unusual. If they are actually so rare, why were they discovered first?

Answer: These were the first exoplanets discovered because they have the strongest gravitational effect on their parent stars, making them easier to detect. (In the same way, cows are much less numerous than mice, but when you look into a pasture, you are more likely to see the cows first.)

Lecture 10

1. Think of several ways that Jupiter's gravity affects asteroids in the solar system.

   Answer: Jupiter perturbs the orbits of the asteroids in the asteroid belt, leading to Kirkwood gaps. These perturbations can also allow asteroids to "collect" around certain orbits, as in the Hilda group and the Trojan asteroids. These are just three of a whole range of gravitational effects due to Jupiter.

2. Imagine a distant binary star system, where a smaller star orbits a larger star. At which of the Lagrange points in this system would you look for planets and asteroids?

   Answer: We should look at the L4 and L5 points of the orbit of the smaller star. The other three Lagrange points (L1 through L3), being unstable, will not easily accumulate natural objects, such as planets and asteroids.

3. The asteroid 3753 Cruithne orbits the Sun in a horseshoe orbit relationship with the Earth. What is the average orbital period of 3753 Cruithne around the Sun? Some have described this asteroid as "Earth's second moon." Do you agree with this description?

   Answer: The average orbital period of 3753 Cruithne is 1 year—a bit less than that when it is in the "catching up" phase and a bit more than that when it is in the "falling behind" phase. Whether you agree with the

description of the asteroid as a moon of the Earth depends on your point of view. On the one hand, like the familiar Moon, the asteroid's orbit around the Sun is intimately linked to that of the Earth. On the other hand, the relation is such that 3753 Cruithne actually stays well away from the Earth; it never comes too close.

Lecture 11

1. Why does the virial theorem not apply to a solid body, such as the Earth? Think about what forces are acting on the particles that make up the Earth.

Answer: A solid body has internal stresses and pressures in addition to gravitational forces. The virial theorem applies only when gravity is the only important force acting on the bodies in the system.

2. A cloud of interstellar gas is just on the verge of contracting due to gravity; that is, its mass is not quite the Jeans mass. Predict what is likely to happen if the cloud gets colder, warmer, larger in radius, or smaller in radius.

Answer: The Jeans mass depends on temperature and density. It increases with temperature and decreases with density. A cloud of a given mass with a higher temperature or a larger radius (lower density) will, therefore, have a larger Jeans mass. Because its actual mass is less than the Jeans mass, it will expand. If the cloud has a lower temperature or a smaller radius (higher density), its Jeans mass will be less, and it will contract.

Lecture 12

1. A rowboat floats in the water. If a person climbs into the rowboat, then the boat sinks a little farther into the water. Explain this in terms of the water pressure on the bottom of the boat.

Answer: The pressure in the water is greater at lower depth. Thus, when the boat sinks a little into the water, the upward pressure force on the

bottom of the boat is greater. This greater force is able to balance the weight of the person.

2. A cork floats in water because of the buoyant force. Would there be any buoyant force if the water and the cork were in an environment without gravity?

   Answer: No. Because there would be no weight force on the water to balance with a pressure force, the pressure in the water would be the same everywhere. The cork would experience no net force.

3. Imagine a chain hanging vertically under its own weight. There is a tension force acting between the links that holds the chain up. Is this force greater at the top of the chain or at the bottom? Compare this situation to hydrostatic equilibrium.

   Answer: The tension force is greatest at the top, because the force must support a greater length of chain. This is just the opposite of the pressure force in hydrostatic equilibrium, which is greatest at the bottom of a body of water. (This makes sense if we think of tension as a kind of "negative pressure.")

Lecture 13

1. Is the field picture of electromagnetic forces (or gravitational forces) fundamentally simpler or more complicated? Explain your answer.

   Answer: Some might say that the field picture is more complicated because it involves both particles and fields. In contrast, the law governing the behavior of a particle is simpler because the particle receives its "marching orders" from the field in its immediate vicinity.

2. In 1905, Einstein laid out two postulates for his special theory of relativity. What would Galileo or Newton have thought of them? Imagine a dialogue among the three of them about the meaning of the principle of relativity.

Answer: Here is how I approached this question: Galileo and Newton would have accepted the first postulate, that every observer sees the same laws of physics. Galileo might have recognized it as a restatement of his own idea. Both would have found the second postulate, about the speed of light, very difficult to understand. Newton in particular stated that space and time have absolute meanings apart from our measurements of length and duration.

Lecture 14

1. You are a producer at a big Hollywood studio. Suppose a visionary director proposes the following: "Instead of merely simulating zero gravity in an airplane—as was done for the movie *Apollo 13*—we should launch a film crew into orbit and make a movie in *actual* zero gravity!" Respond to this (potentially very expensive) movie proposal.

Answer: The occupants of an orbiting spacecraft experience zero gravity for exactly the same reason as the occupants of an airplane in a parabolic arc. It is wrong to say that one is more "authentic" than the other. Thus, this idea does not justify the expense. (The pure publicity value of making a movie in space, however, might be a different matter!)

2. Use the principle of equivalence to show that a horizontal light beam will be deflected downward in a gravitational field. (The bending of light by gravity is the subject of Lecture 18.)

Answer: Consider an accelerating laboratory. A lateral light beam aimed at one spot will miss its target because the lab will increase its speed during the time of flight of the light beam. The beam will appear to be deflected "down" in the apparent gravity of the accelerated lab. By the principle of equivalence, a gravitational field will produce exactly the same effect.

3. Because the tidal effect is tiny over short distances, we may ignore it for our small laboratory. But suppose the lab was larger, and we had to take the tidal effect into account. Would the principle of equivalence still hold as we have stated it?

Answer: No. The principle of equivalence assumes that all parts of the laboratory have the same acceleration. The principle applies, therefore, only if the lab is small enough that any tidal gravitational effect can be ignored. (This is why the tidal effect is the signature of spacetime curvature, as we will see in Lecture 17.)

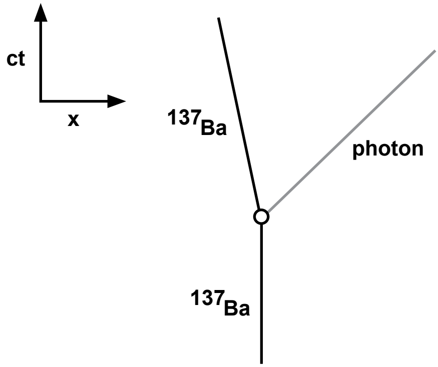
Lecture 15

1. Upon hearing the parable of the surveyors in our lecture, a friend says, "Why should we use coordinates at all? If the surveyors just measured and wrote down all the distances between points in the county, that would contain the same information—and no other surveyors would get different results." How would you answer your friend?

Answer: We could simply record all these distances and give a "coordinate-free" description of geometry. But coordinates are much more efficient. Suppose our map contains 1000 points of interest. Using coordinates, we have to record only 2000 pieces of data (x and y coordinates for each point), but there are $\frac{1}{2}\times 1000\times 999 = 499{,}500$ distances between these points. Thus, the coordinates are a much more "compact" way to describe things.

2. A nucleus of barium-137 is in an "excited" state (that is, a state with extra energy above the minimum). After a couple of minutes, the nucleus emits a gamma ray (a photon of light) and recoils a bit in the opposite direction. Draw a spacetime diagram for this process, showing the nucleus before and after and the photon. If you measure both time and space in m, at what angle should you draw the photon's world line?

Answer: In your spacetime diagram, the original barium nucleus is a vertical world line at the bottom (past) part of the diagram. The decay event is a point. After the decay, the barium world line tilts slightly one way, and the photon world line goes at a 45° angle the other way. The net result is an asymmetrical Y shape.



Lecture 16

1. We defined a straight line on a spherical surface in two ways: as a path of minimum length and as the path followed by someone walking on the surface without turning right or left. Which of these concepts is more like the field idea, in which a moving body gets its marching orders locally? Try to describe great-circle paths on a Mercator map by means of a field that bends and stretches the path.

Answer: The "don't turn right or left" description is more like the field idea because it refers only to a local, step-by-step way to define the path. On the Mercator map, we might imagine that the paths of objects are deflected back toward the equator and accelerated east or west by a field.

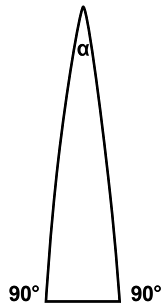
2. Gravitational time dilation, which we introduced in Lecture 14, is extremely small near the Earth's surface. Nevertheless, we now claim that this phenomenon is responsible for the action of gravity on a freely falling object, such as a projectile. How can such a tiny physical cause give rise to such an apparently large effect?

Answer: The answer lies in the fact that the world lines of falling bodies are typically extended very far in the time direction because the speeds involved are quite slow. If I toss an object upward and catch it 1 s later, its *ct* coordinate has changed by 300,000,000 m. From this point of view, the world line only looks very slightly "bent"—an effect that is consistent with the small effect of gravitational time dilation.

Lecture 17

1. We defined a surface of positive curvature as one in which the angles of a triangle add up to more than 180°. We also said that in such a surface, lines that are initially parallel eventually converge. Draw a diagram that shows how these two geometrical facts are connected. (Hint: The parallel lines lie along two sides of a triangle.)

   Answer: My diagram is a long, skinny isosceles triangle. The two base angles are both 90°, so that the other two sides start out parallel. But these sides intersect because the surface has positive curvature; thus, there is a third small angle at the apex. The three angles add up to slightly more than 180°.



2. In the classic science fiction novel *First Men in the Moon*, written in 1901 by H. G. Wells, the scientist Cavor invents a substance called "cavorite" that is "opaque to gravity." Anything placed within a shell of cavorite is no longer affected by the Earth's gravitational pull and, thus, flies in a straight line into space. (This is how the protagonists travel to the Moon.) Comment on cavorite in light of Einstein's general theory of relativity.

   Answer: Cavorite makes no sense in Einstein's theory. Gravity is not something that "radiates from" the Earth or the Sun. Instead, it is a local curvature of spacetime. Because everything moves in the same spacetime, no substance (even cavorite) is immune to this curvature.

Lecture 18

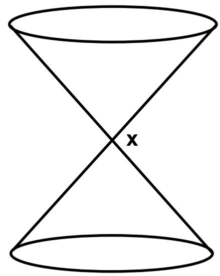
1. It is a remarkable coincidence that our Moon is just the right size to cover the visible disk of the Sun during an eclipse. Could Eddington have made his observations if (a) the Moon were much smaller, or (b) the Moon were much larger?

   Answer: If the Moon were much smaller, it would not block out the light of the Sun, so it would be impossible to see the stars near the

Sun during an eclipse. If the Moon were much larger, the stars nearest the Sun (for which the light-bending effect would be greatest) might be blocked, but some effect could still be seen in the other stars. It is interesting to note that radio telescopes have been used to detect the bending of radio signals near the Sun, even without using eclipses.

2. Every point in spacetime has two light cones associated with it: one that expands outward toward the future and a reversed one that expands outward toward the past. Make a spacetime diagram of this. Events in the forward cone are effectively "in the future," and events in the backward cone are "in the past." Explain this idea. What about events that are in neither light cone?

Answer: Events within the future light cone are those that the event X can affect—the "future." Events within the past light cone can have had an effect on the event X—the "past." Other events can neither cause nor be caused by what happens at X.



Lecture 19

1. A classic museum demonstration of the Earth's rotation is the Foucault pendulum, first exhibited by Leon Foucault in 1851. A large pendulum swings back and forth, but because the Earth is rotating, the plane of the swing slowly precesses around relative to the building. Would you classify this as a gyroscope method or a telescope method?

Answer: Because this method depends on the local behavior of the pendulum, rather than the apparent motion of the distant universe, it counts as a gyroscope method.

2. Suppose a very strong gravitational wave passes through the room in which you are sitting. What sort of damage would you expect?

Answer: The gravitational wave changes the distances between objects. Fluids or flexible structures will move around but suffer little net

damage. Rigid structures, however, such as walls, furniture, and bones, will be subject to tidal stresses. For strong enough gravitational waves, we might expect fractures and warping of these structures.

Lecture 20

1. Some gravitational physicists altogether eschew talking about what happens "inside" a black hole. Instead, they describe the black hole entirely in terms of the properties of its event horizon "surface" and the spacetime outside of it. The inside of a black hole, they say, is not really physical. From a purely philosophical point of view, do they have a point?

Answer: They do. If the interior of a black hole is truly inaccessible, then its exact nature is irrelevant to any experiment or observation we might make and, thus, irrelevant for physics. It might make sense to try to understand the properties of black holes using only the parts of spacetime that are accessible to us.

2. Suppose a miniature black hole of mass 10^{-6} kg formed inside a particle accelerator, and suppose this tiny black hole were initially at rest. Think about what would happen next. (If you like, you might try to estimate the size of the various effects of such a black hole. For example, how large is its event horizon compared to an atom?)

Answer: The event horizon of such a black hole would be so small that it would not be able to swallow an atomic nucleus. The gravitational forces due to the hole would be insignificant compared to the electrostatic attraction between the nucleus and the electron in an atom.

Lecture 21

1. Consider a map of North America. If we consider squares 100 m on a side, the population of each square on the map greatly depends on where we draw the square. If the population in North America were distributed like the mass in the universe, then we should be able to find squares large enough so that the population is more or less the same no

matter where we draw the square. Is the population of North America distributed in this way?

Answer: No. Even if we draw the square 1000 km on a side, it matters very much whether the square is located in a densely populated coastal region or a relatively empty area of the mountainous west or midwest. Thus, the population distribution is not uniform, even on large scales.

2. One objection to Einstein's "static" universe was that it was unstable. Without an exact balance between matter density and the cosmological constant, the cosmos would tend to expand or contract. Why are theories that require this sort of "fine-tuning" less satisfactory than other, more stable theories?

Answer: This is a deep question, but a sensible answer would note that a theory containing special detailed assumptions is less plausible than one that explains the same data without those assumptions. Therefore, we should seek explanations that do not rely on such assumptions.

3. In this lecture, we described how gravitational physics (including cosmology) was in some respects a backwater of physics during the middle part of the 20th century. What determines how important and active a branch of science is at any particular time? What changes and developments would promote a subject to greater importance?

Answer: Many things contribute to the "importance" and "activity" of a scientific field. Some fields play a role in many branches of science; for example, the concept of energy is key to physics, chemistry, and biology. Other scientific fields have important technological applications, such as electromagnetism. Rapid progress is made in a scientific field if that field offers many problems that can be investigated with the available experimental and theoretical techniques. Thus, new discoveries and new techniques can open up a "backwater" field to new exploration. Finally, some scientific fields have a philosophical importance: They tell us something especially significant about the world and our place in it.

Lecture 22

1. When we consider the production of hydrogen and helium in the early universe or the physics of the universe at the time when the cosmic microwave background formed, we do not need to consider the dark energy component of the universe. Why not?

Answer: At these stages in the early universe, the density of matter (including dark matter) was millions of times greater than it is today. The density of dark energy, in contrast, was just the same. Therefore, dark energy was quite insignificant.

2. We think that the cosmic antigravity in the universe remains constant over time. But some physicists have speculated that it might be increasing, growing stronger over time. Suppose that the density of dark energy is actually doubling every few billion years, even as the cosmos expands. What will the far future of our universe be like?

Answer: As the cosmic antigravity becomes more intense, systems that were previously held together by mutual gravity would be pulled apart. First galaxy clusters, then galaxies, and finally star systems, stars, and planets themselves would be pulled to smaller and smaller pieces. Ultimately—many trillions of years hence—even the electromagnetic forces holding atoms together would be overcome. Some physicists call this possible cosmic destiny the "big rip."

Lecture 23

1. Cosmologists sometimes refer to the "dark ages" of the cosmos, the long period of the history of the universe between the creation of the cosmic microwave background (less than 0.5 million years after the big bang) and the formation of the first stars and galaxies (almost 0.5 billion years later). Was anything important happening during this time?

Answer: Yes. The tiny density variations present at the formation of the cosmic microwave background had to be amplified considerably to form

the giant clouds of gas and dark matter that turned into galaxies. This process was driven by gravity.

2. Before the discovery of dark energy, it seemed possible that the universe would expand, halt, and then recollapse after a few hundred billion years. Given that entropy would continue to increase all during that cycle, how would the universe look different during its collapse (the big crunch) than it did after the big bang?

Answer: The big bang was highly uniform—matter was distributed very evenly in space. But gravity has produced a much less uniform distribution of matter. The big crunch, therefore, would be quite "lumpy." The last stages of the collapse would be dominated by giant black holes coalescing in the dense, hot, contracting gas.

Lecture 24

1. At the beginning of our course, we said that the gravitational force had four essential characteristics: (1) Gravity is long range. (2) Gravity is very weak. (3) Gravity is attractive. (4) Gravity is a universal force affecting all masses. How do these observations relate to the revolutionary discoveries of Galileo, Newton, and Einstein?

Answer: Galileo's discovery of the law of free fall was the first indication that gravity affects everything. Newton's mechanics included all four listed properties of gravity. From the universal nature of gravity, Einstein discovered that gravity is not an ordinary force but, rather, an effect of the curvature of spacetime. This insight changes our view of all four properties. In fact, the property "gravity is attractive" is no longer strictly true. The "gravitomagnetic" effects and gravitational waves discussed in Lecture 19 are effects of gravity that have nothing to do with pulling masses closer together.

2. Here's a fun demonstration that you can do at home: Hang a weight from a rubber band. Now heat the rubber band—add entropy to it—using an electric hair dryer. Do you expect the stretched rubber band to get longer or shorter? Make a prediction based on our discussion of polymer

entropy and then perform the experiment to check your prediction. (It helps to mark the original position of the hanging weight in some way to allow you to observe small changes in the length of the rubber band.)

Answer: As you heat up the rubber band, you are adding entropy to it. Therefore, you should expect the rubber band to get shorter (because the polymers have more entropy when they are more tangled up). A careful experiment should confirm this.

3. In this lecture, we suggested that the odd relationship between gravity and entropy might be a clue to the next revolution in gravitational physics. Can you think of other properties of gravity that might suggest other possibilities?

Answer: Your responses to this speculative question may vary. Maybe you will be the one to guess the key to the next revolution! One interesting thought: Why is gravity always an attractive force? Does Einstein's theory explain this in a natural way? Or is there a deeper view?

Answer Key—Problems

Lecture 2

Problem: How important was air resistance in Felix Baumgartner's long skydive from 39,000 m? Here is an interesting calculation: Show that, at a constant acceleration of 10 m/s^2, Baumgartner would fall to the ground in less than 90 seconds. (You should use Galileo's relation between distance, acceleration, and time.)

Answer: In 90 seconds, Baumgartner would fall

$d = \frac{1}{2}\left(10\frac{\text{m}}{\text{s}^2}\right)(90\text{ s})^2 = 40{,}500$ m, a greater distance than his altitude.

Lecture 3

Problem: The semimajor axis of Saturn's orbit around the Sun is 9.6 AU. Find the length of Saturn's year (i.e., its orbital period about the Sun).

Answer: Using units of AUs and Earth years, Kepler's harmonic law states $P^2 = a^3$. Thus, $P = \sqrt{9.6^3} = 29.7$ years.

Lecture 4

Problem: The mass of the Earth is about 6×10^{24} kg, and the radius of the Earth is 6.4×10^6 m (about 6400 km). From this information and the value of Newton's constant, G (6.67×10^{-11} m^3/kg s^2), calculate the acceleration due to gravity near the Earth's surface. Does the value you obtain agree with the value we gave in the last lecture (about 10 m/s^2)?

Answer: $a = \frac{GM}{R^2} = \frac{\left(6.67\times10^{-11}\right)\left(6.0\times10^{24}\right)}{\left(6.4\times10^{6}\right)^2} = 9.8\text{ m/s}^2,$

which agrees closely.

Lecture 5

Problem: Use the Newtonian equations from the previous lectures to calculate the acceleration of gravity due to a 7.25-kg shot put at a distance of 0.10 m from its center. At this acceleration, how far would an object "fall" in 1000 seconds (a bit more than a quarter of an hour)?

Answer: From Lecture 4, the acceleration is $\frac{GM}{R^2} = \frac{\left(6.67\times10^{-11}\right)\left(7.25\right)}{\left(0.10\right)^2} = 4.8\times10^{-8}$ m/s^2. From Lecture 2, the distance traveled would be $d = \frac{1}{2}at^2 = \frac{1}{2}\left(4.8\times10^{-8}\right)\left(1000\right)^2 = 0.024$ m, which is about 1 inch.

Lecture 7

Problem: In this lecture, we saw that the speed of a body in a circular orbit of radius r around a planet or star of mass M is $v_{circ} = \sqrt{\frac{GM}{r}}$. In each orbit, the satellite traverses a total distance of $d = 2\pi r$. A body moving with a speed v travels a distance d in a time $t = \frac{d}{v}$. From these pieces of information, derive the harmonic law for circular orbits.

Answer: The orbital period is $T^2 = \left(\frac{2\pi d}{v_{circ}}\right)^2 = \frac{4\pi^2}{G}\frac{r^3}{M}$. Since the semimajor axis of a circular orbit is equal to r, this is Newton's form of Kepler's harmonic law: $\left(\frac{GM}{4\pi^2}\right)T^2 = a^3$.

Lecture 8

Problem: In the science fiction story "Neutron Star" by Larry Niven, the hero, Beowulf Shaeffer, flies his spaceship just 2000 km (2×10^6 m) from a neutron star, a very compact body about the mass of the Sun (2×10^{30} kg). At the risk of spoiling the story, Shaeffer has a lot of trouble with the tidal effect! Calculate the relative tidal acceleration of two apples placed at Shaeffer's head and feet, which are 2 m apart.

Answer: $a = \frac{GM}{r^2}\left(\frac{2h}{r}\right) = \frac{\left(6.67\times10^{-11}\right)\left(2\times10^{30}\right)}{\left(2\times10^{6}\right)^2}\frac{2(2)}{2\times10^6} = 67\ \text{m/s}^2$, which is more than 6 times the acceleration of gravity at the Earth's surface.

Lecture 11

Problem: The virial theorem says that PE = –2KE for a galaxy cluster. Suppose only 1/3 of the mass of a galaxy cluster is visible. If we write PE′ = –κKE′ for the visible potential and kinetic energies, what is the coefficient κ? Is it greater or less than 2? If we include only the visible matter, does the total energy, KE′ + PE′, appear to be positive or negative?

Answer: $\text{KE}' = \frac{1}{3}\text{KE}$, since KE is proportional to the mass.

$\text{PE}' = \frac{1}{9}\text{PE}$, since PE is proportional to the mass squared.

Thus, $\text{PE}' = \frac{1}{9}\text{PE} = -\frac{2}{9}\text{KE} = -\frac{6}{9}\text{KE}'$,

and so $\kappa = 2/3$, which is smaller than 2. The total energy appears to be positive.

Lecture 15

Problem: A moving clock emits two flashes of light. When the clock is seen at rest, the two flashes are emitted just 12 m of time (40 ns) apart. Now we observe the same clock moving so that the two flashes occur 5 m apart in space. How much time (in m and ns) separates the two flashes in this frame of reference? By what factor is the moving clock running slowly?

Answer: In the first frame, the interval s = 12 m. In the second frame,

$s^2 = \left(c\Delta t\right)^2 - \Delta x^2$, so $c\Delta t = \sqrt{s^2 + \Delta x^2} = \sqrt{12^2 + 5^2} = 13$ m, or about 43.3 ns.

The clock is running 13/12 = 1.08 times slower.

Lecture 17

Problem: In the lecture, we said that we could derive a version of Einstein's equation from what we already know. Do this by imagining six apples on the top, bottom, forward, back, left, and right sides of a sphere of radius r that contains mass of density ρ . Here are the equations you need:

- Acceleration of each apple toward center of sphere: $a = \frac{GM}{r^2}$
- Mass of sphere: $M = V\rho$
- Volume of sphere: $V = \frac{4\pi}{3}r^3$
- Relation of curvature to relative acceleration: $a = \mathcal{R}c^2d$

Answer: The separation of each pair of apples is $d = 2r$. Each one accelerates toward the other, so the relative acceleration is positive and equals twice the gravitational acceleration of each one. For each pair of apples: $\mathcal{R} = \frac{a}{c^2d} = \frac{1}{c^2\left(2r\right)}\left(\frac{2G\rho V}{r^2}\right) = \frac{G\rho}{c^2r^3}\left(\frac{4\pi}{3}r^3\right) = \frac{4\pi G}{3c^2}\rho$

Since this is true for each of the three pairs of apples, the total curvature is $\mathcal{R}_{\text{total}} = \frac{4\pi G}{c^2}\rho$, as stated in the lecture.

Lecture 20

Problem: Calculate the radius of black holes with the mass of the Earth (6×10^{24} kg) and a galaxy of 10^{11} times the mass of the Sun.

Answer: For the Earth-mass black hole,

$$R_S = \frac{2GM}{c^2} = \frac{2\left(6.67\times10^{-11}\right)\left(6\times10^{24}\right)}{\left(3\times10^{8}\right)^2} = 0.0089 \text{ m, just 9 mm.}$$

For the galaxy-mass black hole,

$$R_S = \frac{2GM}{c^2} = \frac{2\left(6.67\times10^{-11}\right)\left(10^{11}\times2\times10^{30}\right)}{\left(3\times10^{8}\right)^2} = 2.96\times10^{14} \text{ m, or 300 billion km.}$$

Lecture 24

Problem: If an atomic nucleus (radius = 10^{-15} m) were magnified to the size of a planet (radius = 10^{7} m), would the Planck length be larger or smaller than the original nucleus?

Answer: The nucleus is expanded 10^{22} times. This means that the new Planck length would be $\left(\text{new } L_P\right) = 10^{22}\left(\text{actual } L_P\right) = 4.0\times10^{-13}$ m, which is about 400 times larger than the original nuclear radius (but still much smaller than an atom).

Bibliography

Baez, John. "Lagrange Points." http://math.ucr.edu/home/baez/lagrange.html. John Baez is a mathematical physicist at the University of California, Riverside. He is also a compulsive and prolific explainer of scientific ideas that catch his interest. His web page on Lagrange points, with many links and references, is a fine place to begin a deeper look at these remarkable orbital patterns.

Beatty, J. Kelly, Carolyn Collins Peterson, and Andrew Chaikin, eds. *The New Solar System*. 4th ed. Cambridge: Cambridge University Press, 1999. This series of interconnected articles by experts in planetary science provides a vivid snapshot of our contemporary understanding of the planetary system we inhabit. The effects of gravity, via tides and orbital perturbations, play a central role in many chapters, including chapter 16 ("Planetary Rings") and chapter 25 ("Asteroids").

Case Western Reserve University, "Galaxy Crash JavaLab!" http://burro.cwru.edu/JavaLab/GalCrashWeb/. On this site, you can perform your own simulations of galaxy collisions with a gravitational *N*-body program. The simulated galaxies automatically obey the virial theorem.

Collins, Harry. *Gravity's Shadow: The Search for Gravitational Waves*. Chicago: University of Chicago Press, 2004. A fascinating and detailed account of the 40-year struggle to detect gravitational waves. Collins describes the theory and the experiments, but he is even more interested in the way science works in developing a new field.

Dolling, Lisa M., Arthur F. Gianelli, Glenn N. Statile, eds. *The Tests of Time: Readings in the Development of Physical Theory*. Princeton: Princeton University Press, 2003. The creators of this book have done a great service by selecting and assembling dozens of great readings from primary sources in the history of physics. Newton explains the basic facts about gravitation starting on page 115; Eddington announces the bending of starlight on page

330; Hubble describes the expanding universe on page 602. I use this book as a text in one of my own seminar courses.

Doody, Dave. *Basics of Space Flight*. Pasadena, CA: Blüroof Press, 2011. Dave Doody is an engineer at NASA's Jet Propulsion Laboratory and has been involved in spacecraft operations for more than two decades. This book is adapted from a popular series of articles (and a class offered at JPL), aiming to present to the general public the essential ideas and techniques used in actual space flight. The first third of the book is all about gravity and orbital mechanics.

Einstein, Albert, and Leopold Infeld. *The Evolution of Physics*. New York: Touchstone, 1966. First published in 1938, this brief, popular book lets us look at the development of physics through the eyes of one of the heroes of gravitational physics and one of his key coworkers. Einstein and Infeld lucidly survey how the mechanical universe of Galileo and Newton gave way to the field concept of Faraday and Maxwell, leading finally to the revolutions of relativity and quantum theory.

Gates, Evalyn. *Einstein's Telescope: The Hunt for Dark Matter and Dark Energy in the Universe*. New York: W.W. Norton, 2009. Gates tells a lively and up-to-the-minute story about gravitational lensing and its use to map dark matter in the cosmos. The author is now the executive director of the Cleveland Museum of Natural History, but in her other life as a well-known astrophysicist and cosmologist, she played an important role in the very discoveries she describes.

Feynman, Richard Phillips. *The Character of Physical Law*. Cambridge: MIT Press, 1965. Richard Feynman's 1964 "Messenger Lectures" at Cornell University are a brilliant presentation of the scope and meaning of physics. He begins with Newton's law of universal gravitation as a prototype of physical law.

Galileo Galilei. *Dialogue Concerning the Two Chief World Systems*. New York: Modern Library, 2001. This classic work is presented as a three-way conversation among imaginary characters, exploring the astronomical theories of Ptolemy and Copernicus. (This is my favorite translation, that

of renowned Galileo scholar Stillman Drake.) Here, Galileo presents the fruits of a lifetime of thought and discovery about astronomy and mechanics. His description of the principle of relativity (in the "Second Day" of the dialogue) has never been surpassed for clarity and insight. His theory of the tides (in the "Fourth Day") is probably Galileo's greatest blunder. It was this book—a little too convincing and accessible to the lay reader—that spurred the ecclesiastical authorities of the day to take action against Galileo.

Gamow, George. *Gravity*. Toronto: Dover, 2003. Gamow, a pioneer in both quantum mechanics and cosmology, was also a remarkable writer of popular science books. This brief book, originally written in 1962, does not shy away from doing a little mathematics where it is helpful, but the explanations are always enlivened by the author's appealing prose and charming cartoons (for which the author claims equal inspiration from the Sunday newspaper comics and the Italian painter Botticelli).

Hawking, Stephen, ed. *On the Shoulders of Giants*. Philadelphia: Running Press, 2002. Famed gravitational physicist Stephen Hawking has brought together in one volume some of the classic works by Copernicus, Galileo, Kepler, Newton, and Einstein, along with introductions and explanations by Hawking himself. This book belongs on the shelf of every student of the history of physics.

Lorentz, H. A., A. Einstein, H. Minkowski, and H. Weyl, with notes by A. Sommerfeld. *The Principle of Relativity: A Collection of Original Papers on the Special and General Theory of Relativity*. Toronto: Dover, 1952. This collection of the original scientific papers on relativity by Einstein and his contemporaries contains some real gems. Especially rewarding is Hermann Minkowski's famous lecture "Space and Time," giving a new geometrical view of Einstein's theory.

Moffat, Peter. *Einstein and Eddington*. Directed by Philip Martin. BBC Films, 2008. This 90-minute film, originally produced for broadcast in the United Kingdom, shows how the development and acceptance of general relativity came about during the era of the First World War. Einstein and Eddington were both pacifists whose personal nonconformism went well beyond their scientific pursuits. The script somewhat compresses and overdramatizes the

historical story, but the performances—by David Tenant (of *Doctor Who* fame), who plays Eddington, and Andy Serkis (Gollum from *The Lord of the Rings*), who plays Einstein—make this film well worth watching.

Schumacher, Benjamin. *The Pasadena Rule*. http://www.raygunrevival.com/ (issues 29, 31, 33, 35, 37 and 39). My sky-diving science fiction novella, published as a six-part serial in the online magazine *Ray Gun Revival*. Also available as an e-book from amazon.com.

Smolin, Lee. *Three Roads to Quantum Gravity*. London: Basic Books, 2002. Lee Smolin is an iconoclastic theoretical physicist with a distinctive (and controversial) point of view, but he is also a clear writer about the deepest problems facing gravitational physics. In this book, he surveys some of the best current thinking about how to marry quantum mechanics and general relativity. Could it be that these apparently disparate approaches could all be shadows of a single, deeper insight?

Stukeley, William. "Newton's Apple." http://royalsociety.org/library/moments/newton-apple/. William Stukeley's version of the story about Isaac Newton and the apple can be read online, courtesy of the Royal Society.

Taylor, Edwin F., and John A. Wheeler. *Spacetime Physics*. New York: W. H. Freeman, 1992. A beautiful introduction to relativity, both special and general, from the spacetime point of view. Though it makes some moderate mathematical demands on the reader, this book also presents the essential physics in a remarkable and original way. The numerous problems and exercises are some of the best parts of the book, and full solutions are given for all of them.

Thorne, Kip S. *Black Holes and Time Warps: Einstein's Outrageous Legacy*. New York: W.W. Norton, 1994. Though he is a (now retired) professor at Caltech, Thorne gives a very accessible account of physics in curved spacetime, including black holes, Hawking radiation, and more. Here, you will find some of the most exciting developments in contemporary physics, told by one of the key participants.

Tyson, Neil deGrasse. *Death by Black Hole and Other Cosmic Quandaries*. New York: W.W. Norton, 2007. A collection of brief essays, first written for *Natural History* magazine, by one of the best contemporary explainers of science. Tyson not only gets into the gruesome details of what would happen to you if you fell into a black hole, but he also discusses Lagrange points and many other aspects of gravitational physics.

University of St. Andrews, School of Mathematics and Statistics. "Online Dictionary of Mathematical Biography." http://www-history.mcs.st-and.ac.uk/BiogIndex.html. This invaluable website provides in-depth biographical essays for many mathematicians and mathematical physicists, including many of those discussed in our course. The wider "MacTutor History of Mathematics Archive" is also on the same site; together, the two form an invaluable resource for anyone exploring the history of science. For instance, a useful survey of pre-Newtonian theories of gravity can be found here: http://www-history.mcs.st-andrews.ac.uk/HistTopics/Gravitation.html.

Verlinde, Erik P. "On the Origin of Gravity and the Laws of Newton." http://arxiv.org/abs/1001.0785. Though some of the discussion (e.g., the AdS/CFT correspondence) will seem a bit arcane to the nonspecialist, Verlinde's recent paper is actually fairly readable and well worth perusal.

Wald, Robert M. *Space, Time and Gravity: The Theory of the Big Bang and Black Holes*. 2nd ed. Chicago: University of Chicago Press, 1992. With a little mathematics and a lot of physical insight, Wald (a well-known gravitational physicist at the University of Chicago) presents relativity, basic cosmology, and black hole physics in a compact and readable volume. Wald aims his survey at a serious lay reader, but even a specialist will find much value in this book.

Wheeler, John A. *A Journey into Gravity and Spacetime*. New York: Scientific American Library, 1990. A fascinating and unique view of general relativity from one of its great modern practitioners. Wheeler summarized Einstein's theory this way: "Spacetime tells matter how to move; matter tells spacetime how to curve." This book is a playful and often surprising explication of that essential idea.